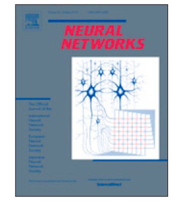
Neural Networks 164 (2023) 588–605

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet



Design of continuous-time recurrent neural networks with piecewise-linear activation function for generation of prescribed sequences of bipolar vectors

Norikazu Takahashi [a,*], Tsuyoshi Yamakawa [b], Yasuhiro Minetoma [b], Tetsuo Nishi [b], Tsuyoshi Migita [a]

[a] *Okayama University, 3–1–1 Tsuhima-naka, Kita-ku, Okayama, 700-8530, Japan*
[b] *Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, 819-0395, Japan*

A R T I C L E   I N F O

Article history:
Received 10 September 2022
Received in revised form 16 March 2023
Accepted 9 May 2023
Available online 13 May 2023

Keywords:
Recurrent neural network
Piecewise-linear activation function
Sequence
Bipolar vector
Mathematical programming
Limit cycle

A B S T R A C T

A recurrent neural network (RNN) can generate a sequence of patterns as the temporal evolution of the output vector. This paper focuses on a continuous-time RNN model with a piecewise-linear activation function that has neither external inputs nor hidden neurons, and studies the problem of finding the parameters of the model so that it generates a given sequence of bipolar vectors. First, a sufficient condition for the model to generate the desired sequence is derived, which is expressed as a system of linear inequalities in the parameters. Next, three approaches to finding solutions of the system of linear inequalities are proposed: One is formulated as a convex quadratic programming problem and others are linear programming problems. Then, two types of sequences of bipolar vectors that can be generated by the model are presented. Finally, the case where the model generates a periodic sequence of bipolar vectors is considered, and a sufficient condition for the trajectory of the state vector to converge to a limit cycle is provided.

1. Introduction

Recurrent neural networks (RNNs) (Aihara et al., 1990; Amari, 1972; Chua & Yang, 1988a; Elman, 1990; Hochreiter & Schmidhuber, 1997; Hopfield, 1982, 1984; Jordan, 1997) have been extensively studied in the past decades due to their important applications in various areas such as associative memory (Amari, 1972; Hopfield, 1982, 1984; Michel & Farrell, 1990), combinatorial optimization (Hertrich & Skutella, 2021; Hopfield & Tank, 1986), image and video processing (Chua & Yang, 1988b; Pan et al., 2016; Takahashi et al., 2010), acoustic modeling (Sak et al., 2014), natural language processing (Lawrence et al., 2000; Lee & Dernoncourt, 2016) and time series prediction (Hewamalage et al., 2021). Unlike feedforward neural networks, RNNs have feedback connections between neurons. Hence the output of each neuron depends not only on the current input but also on the internal state. This property allows an RNN to exhibit temporal dynamic behavior and to generate a sequence of patterns as the temporal evolution of the outputs of the neurons. For more details on the history and recent advances in RNNs, the readers are referred to some survey papers (Lipton et al., 2015; Salehinejad et al., 2017; Yu et al., 2019).

As with feedforward neural networks, the universality of RNNs has been theoretically proved from various perspectives. Siegelmann et al. proved that there exists a discrete-time RNN with hidden neurons and external inputs that simulates an arbitrary Turing machine (Kilian & Siegelmann, 1996; Siegelmann & Sontag, 1991, 1994, 1995). Hammer proved that any measurable function from lists of real vectors to a real vector space can be approximated arbitrarily well in probability by an RNN with hidden neurons (Hammer, 2000). Schäfer and Zimmermann proved that open dynamical systems can be approximated by a discrete-time RNN in state space model having external inputs with an arbitrary accuracy (Schäfer & Zimmermann, 2007). This result was recently extended to the case of stochastic inputs (Chen et al., 2022). Funahashi and Nakamura proved that any finite time trajectory of a given continuous-time dynamical system can be approximately realized by a continuous-time RNN with some hidden neurons (Funahashi & Nakamura, 1993). Takahashi et al. proved that all of the 256 possible local pattern sets can be realized by a one-dimensional two-layer cellular neural network model (Takahashi et al., 2008). Li et al. recently studied the approximation properties and optimization dynamics of continuous-time linear RNNs when applied to learn input–output relationships in temporal data (Li et al., 2022).

Although RNNs have been proved to be universal, finding the parameters of an RNN so that it generates given sequences of

* Corresponding author.
E-mail address: takahashi@okayama-u.ac.jp (N. Takahashi).

https://doi.org/10.1016/j.neunet.2023.05.013

patterns is not an easy task. Hence a large number of studies related to this issue have been conducted. Amari studied the stability of discrete-time RNNs with the sign activation function, and presented a condition for such an RNN with parameters determined from some sequences of bipolar vectors using Hebbian rule to generate each of the sequences (Amari, 1972). Mori et al. considered the same type of RNNs as Amari, and evaluated through numerical experiments their sequence retrieval capability when the parameters are determined from some periodic sequences of bipolar vectors using Hebbian rule (Mori et al., 1989). Williams and Zipser considered a class of discrete-time RNNs with external inputs, and proposed a real-time learning algorithm for such an RNN based on the gradient descent on the error defined as the difference between the target outputs and the actual outputs of the RNN at each time step (Williams & Zipser, 1989). Pearlmutter considered a class of continuous-time RNNs with external inputs, and proposed a learning method for such an RNN based on the gradient descent on the error defined as the difference between the target trajectory and the state trajectory of the RNN for a given period of time (Pearlmutter, 1989). Yoneyama et al. proposed a linear programming-based design method for a class of continuous-time RNNs with the sign activation function to generate one or more periodic sequences of bipolar vectors, and demonstrated that an RNN generating two given periodic sequences of four dimensional bipolar vectors can be obtained (Yoneyama et al., 2001).

Despite considerable efforts by many researchers, it is still not fully understood what kind of sequences can be generated by RNNs having neither external inputs nor hidden neurons. Finding answers to this question is an important step not only to gain a deeper understanding of the fundamental properties of RNNs but also to take full advantage of the potential capabilities of RNNs for various applications.

In this paper, we focus our attention on a continuous-time RNN model, which uses the piecewise-linear activation function $f(x) = (|x+1| - |x-1|)/2$ and has neither external inputs nor hidden neurons, and study the problem of finding the parameters of the model so that it generates a given sequence of bipolar vectors. Here, the term "bipolar" means that each entry of the vector is either $+1$ or -1. In order to make the problem more precise, we need to discuss two issues. One is the gap between the desired sequence of bipolar vectors and the actual output vector of the RNN. Since the model is described by a system of ordinary differential equations, the output vector varies continuously with time, that is, it cannot jump from one bipolar vector to another. We thus assume that two consecutive bipolar vectors in the desired sequence differ in one and only one entry, and allow the output vector of the RNN to take intermediate values between those two bipolar vectors. The other issue is the correspondence between the output and state vectors of the RNN. A bipolar output vector does not determine a unique state vector, due to the saturation property of the activation function. We thus need to consider all state vectors corresponding to the first bipolar vector in the desired sequence as the initial condition.

The contributions of this paper are summarized as follows. First, we derive a sufficient condition for the model to generate a given sequence of bipolar vectors, which is expressed as a system of linear inequalities in the parameters. Second, we propose three approaches to finding the parameters with certain properties while satisfying the system of linear inequalities: one is formulated as a convex quadratic programming problem and the others are linear programming problems. The differences among these three approaches are demonstrated through an example. Third, we present two types of sequences of bipolar vectors that can be generated by the model. In one type, the sequence starts with the vector of all 1's, replaces 1 with -1 one by one in order from the first entry to the last one until the vector of all -1's is obtained, and replaces -1 with 1 one by one in the same order until the initial vector is obtained. In the other type, the sequence starts with the vector of all 1's, replaces 1 in the first entry with -1, moves -1 toward the last entry, and returns to the initial vector. Finally, we consider the case where the model generates a periodic sequence of bipolar vectors, and provide a sufficient condition for the trajectories of the state vector to converge to a limit cycle.

The RNN design approaches proposed in this paper have potential applications in associative memory, binary image processing and central pattern generators, for example. RNNs designed so that they generate multiple sequences of bipolar vectors may be applicable to auto- or hetero-associative memory. As for image processing, it is well known that a special class of the RNN model considered in this paper can perform various image processing tasks (Chua & Yang, 1988b; Takahashi et al., 2010). So the proposed approaches may expand the range of applications in image processing. RNNs designed so that they generate a periodic sequence of bipolar vectors may be used as central pattern generators for robot locomotion (Ijspeert, 2008), because the trajectories of the state vector of such an RNN are likely to converge to a limit cycle.

In the previous studies on nonlinear dynamics of continuous-time RNNs, the main focus was on the convergence of the trajectories of the state vector to equilibrium points (Liu et al., 2021). For example, for the RNN model considered in this paper, it is known that if the weight matrix satisfies a certain condition then the trajectories of the state vector converges to one of the equilibrium points for any initial condition (Takahashi & Chua, 1998). Also, for RNN models with different types of piecewise-linear activation functions, stability and instability of multiple equilibrium points have been investigated (Deng et al., 2023; Nie & Zheng, 2015a, 2015b, 2016). In contrast, this paper studies the convergence of the trajectories of the state vector to a limit cycle. In particular, it is shown through some examples that we can design RNNs having a stable limit cycle using the methods developed in this paper. Generation of stable limit cycles is a fundamental research topic in the field of control systems (Azhdari & Binazadeh, 2021; Benmiloud et al., 2018). The results of this paper may provide a new direction in this field.

Preliminary results of this work were presented in two international conference papers (Takahashi & Minetoma, 2008; Takahashi et al., 2005). However, these papers consider only the case where the desired sequence is periodic, that is, the last bipolar vector in the sequence is equal to the first one. Hence the results in Section 3 are slightly different from the ones in those papers. Also, the sufficient condition for the convergence of the state trajectories to a limit cycle presented in Section 6 is different from the one in the paper (Takahashi & Minetoma, 2008) because the latter requires infinite number of steps to test. In addition, the development of the three design methods in Section 4 and the discovery of the realizable sequences of bipolar vectors in Section 5 are original contributions of this paper.

Throughout this paper, we use the following notations. The set of real numbers is denoted by $\mathbb{R}$. The set of positive integers is denoted by $\mathbb{Z}_{++}$. For a function g of time t, the time derivative is denoted by $\dot{g}(t)$. Square brackets are used to form vectors and matrices. Also, comma separated lists of scalars enclosed with parentheses are used to form column vectors. Therefore, by $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, we mean that

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1 \; x_2 \; \cdots \; x_n]^{\mathrm{T}}$$

where the superscript T denotes transposition. For a square matrix $\boldsymbol{J}$, the matrix norm induced by ℓ_2-norm or the Euclidean norm of vectors is denoted by $\|\boldsymbol{J}\|_2$. The natural logarithm is denoted by ln. For a subset M of $\mathbb{R}^n$, the boundary of M is denoted by ∂M.

2. Recurrent neural network model and problem statement

We consider a continuous-time recurrent neural network (RNN) model described by the system of differential equations:

$$\dot{x}_i(t) = -x_i(t) + \sum_{j=1}^{n} w_{ij} f(x_j(t)) + b_i,\ i = 1, 2, \ldots, n \tag{1}$$

with

$$f(x) = \frac{1}{2}(|x+1| - |x-1|), \tag{2}$$

where $x_i(t) \in \mathbb{R}$ and $f(x_i(t)) \in [-1, 1]$ represent the state and output of neuron i at time t respectively, $\dot{x}_i(t)$ denotes the time derivative of $x_i(t)$, $w_{ij} \in \mathbb{R}$ is the weight of the connection from neuron j to neuron i, $b_i \in \mathbb{R}$ is the bias of neuron i. Throughout this paper, we assume that

$$w_{ii} = 2,\ i = 1, 2, \ldots, n. \tag{3}$$

Eq. (1) can be rewritten in vector form as

$$\dot{\boldsymbol{x}}(t) = -\boldsymbol{x}(t) + \boldsymbol{W}\boldsymbol{f}(\boldsymbol{x}(t)) + \boldsymbol{b}$$

where

$$\boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix},\quad \boldsymbol{f}(\boldsymbol{x}(t)) = \begin{bmatrix} f(x_1(t)) \\ f(x_2(t)) \\ \vdots \\ f(x_n(t)) \end{bmatrix},$$

$$\boldsymbol{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix},\quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

In what follows, $\boldsymbol{x}(t)$ is called the state vector, $\boldsymbol{f}(\boldsymbol{x}(t))$ the output vector, $\boldsymbol{W}$ the weight matrix, and $\boldsymbol{b}$ the bias vector. The trajectory of the state vector $\boldsymbol{x}(t)$ passing through a point $\boldsymbol{x}^0 \in \mathbb{R}^n$ at $t = 0$ is denoted by $\boldsymbol{\phi}(t, \boldsymbol{x}^0) = (\phi_1(t, \boldsymbol{x}^0), \phi_2(t, \boldsymbol{x}^0), \ldots, \phi_n(t, \boldsymbol{x}^0))$. Since f is Lipschitz continuous with the Lipschitz constant 1, the existence and uniqueness of $\boldsymbol{\phi}(t, \boldsymbol{x}^0)$ are guaranteed by the Picard–Lindelöf theorem.

The problem we consider in this paper is, roughly speaking, to find $\boldsymbol{W}$ and $\boldsymbol{b}$ of the RNN model so that it generates a given sequence of bipolar vectors as the temporal evolution of the output vector. For simplicity, we assume that two consecutive bipolar vectors in the sequence differ in one and only one entry. Also, since all trajectories of the output vector of the RNN model is continuous, we allow the output vector to take intermediate values during the transition from one bipolar vector to another. Then the problem is formally stated as follows.

Problem 1. Given a sequence $S = \{\boldsymbol{\alpha}^0, \boldsymbol{\alpha}^1, \ldots, \boldsymbol{\alpha}^m\}$ of n-dimensional bipolar vectors such that

1. $\boldsymbol{\alpha}^{k+1}$ differs from $\boldsymbol{\alpha}^k$ in one and only one entry for $k = 0, 1, \ldots, m-1$, and
2. $\boldsymbol{\alpha}^k = \boldsymbol{\alpha}^{k'}$ holds only if $k = 0$ and $k' = m$,

find the weight matrix $\boldsymbol{W}$ and the bias vector $\boldsymbol{b}$ of the RNN model described by (1)–(3) so that if $\boldsymbol{f}(\boldsymbol{x}(0)) = \boldsymbol{\alpha}^0$ then there exists a sequence $0 = t^{0-} \le t^{0+} < t^{1-} \le t^{1+} < \cdots < t^{m-} < \infty$ such that $\boldsymbol{f}(\boldsymbol{x}(t)) = \boldsymbol{\alpha}^k$ $(t^{k-} \le t \le t^{k+})$ and $\boldsymbol{f}(\boldsymbol{x}(t)) \in \{\theta\boldsymbol{\alpha}^k + (1-\theta)\boldsymbol{\alpha}^{k+1} \mid 0 < \theta < 1\}$ $(t^{k+} < t < t^{(k+1)-})$ for $k = 0, 1, \ldots, m-1$ and $\boldsymbol{f}(\boldsymbol{x}(t^{m-})) = \boldsymbol{\alpha}^m$.

Let us consider the case where $S = \{(1, 1), (1, -1), (-1, -1), (-1, 1), (1, 1)\}$. It is clear that this sequence satisfies the conditions in Problem 1. In this case, we have to find $\boldsymbol{W}$ and $\boldsymbol{b}$ of the RNN model so that if $\boldsymbol{f}(\boldsymbol{x}(0)) = (1, 1)$ then there exists a sequence $0 \le t^{0+} < t^{1-} \le t^{1+} < \cdots < t^{4-}$ such that $\boldsymbol{f}(\boldsymbol{x}(t)) = (1, 1)$ $(0 \le t \le t^{0+})$, $f(x_1(t)) = 1$ and $|f(x_2(t))| < 1$ $(t^{0+} < t < t^{1-})$, $\boldsymbol{f}(\boldsymbol{x}(t)) = (1, -1)$ $(t^{1-} \le t \le t^{1+})$, $|f(x_1(t))| < 1$ and $f(x_2(t)) = -1$ $(t^{1+} < t < t^{2-})$, $\boldsymbol{f}(\boldsymbol{x}(t)) = (-1, -1)$ $(t^{2-} \le t \le t^{2+})$, $f(x_1(t)) = -1$ and $|f(x_2(t))| < 1$ $(t^{2+} < t < t^{3-})$, $\boldsymbol{f}(\boldsymbol{x}(t)) = (-1, 1)$ $(t^{3-} \le t \le t^{3+})$, $|f(x_1(t))| < 1$ and $f(x_2(t)) = 1$ $(t^{3+} < t < t^{4-})$, and $\boldsymbol{f}(\boldsymbol{x}(t^{4-})) = (1, 1)$.

It should be noted that the initial value $\boldsymbol{x}(0)$ of the state vector is not uniquely determined from the assumption $\boldsymbol{f}(\boldsymbol{x}(0)) = \boldsymbol{\alpha}^0$. In fact, the assumption holds if and only if $x_i(0)\alpha_i^0 \ge 1$ for $i = 1, 2, \ldots, n$. We thus formulate the problem more precisely in terms of the trajectories of the state vector. Let the intervals $(-\infty, -1]$, $(-1, 1)$ and $[1, \infty)$ be denoted by I_{-1}, I_0 and I_1, respectively. For any n-dimensional vector $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ with $v_i \in \{-1, 0, 1\}$ for all i, let the region $\mathcal{R}_{\boldsymbol{v}} \subset \mathbb{R}^n$ be defined as

$$\mathcal{R}_{\boldsymbol{v}} := \left\{(x_1, x_2, \ldots, x_n) \mid x_i \in I_{v_i}, i = 1, 2, \ldots, n\right\}. \tag{4}$$

Using these notations, the problem is stated in a more precise way as follows.

Problem 2. Given a sequence $S = \{\boldsymbol{\alpha}^0, \boldsymbol{\alpha}^1, \ldots, \boldsymbol{\alpha}^m\}$ of n-dimensional bipolar vectors satisfying the two conditions in Problem 1, find the weight matrix $\boldsymbol{W}$ and the bias vector $\boldsymbol{b}$ of the RNN model described by (1)–(3) so that for any $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}^0}$ the trajectory $\boldsymbol{\phi}(t, \boldsymbol{x}^0)$ of the state vector visits $\mathcal{R}_{(\boldsymbol{\alpha}^0+\boldsymbol{\alpha}^1)/2}$, $\mathcal{R}_{\boldsymbol{\alpha}^1}$, $\mathcal{R}_{(\boldsymbol{\alpha}^1+\boldsymbol{\alpha}^2)/2}, \ldots, \mathcal{R}_{\boldsymbol{\alpha}^m}$ in this order.

In what follows, we say that the RNN model generates the sequence S of bipolar vectors in Problem 2 if the trajectories of the state vector behave as described in Problem 2.

Let us consider, for example, a two-neuron RNN having the following weight matrix and the bias vectors:

$$\boldsymbol{W} = \begin{bmatrix} 2.0 & -1.6 \\ 1.7 & 2.0 \end{bmatrix},\quad \boldsymbol{b} = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}. \tag{5}$$

Fig. 1 shows three trajectories $\boldsymbol{\phi}(t, (1, 1))$, $\boldsymbol{\phi}(t, (5, 1))$ and $\boldsymbol{\phi}(t, (1, 5))$ of the state vector generated by this RNN. It is seen that these trajectories visit $\mathcal{R}_{(0,1)}$, $\mathcal{R}_{(-1,1)}$, $\mathcal{R}_{(-1,0)}$, $\mathcal{R}_{(-1,-1)}$, $\mathcal{R}_{(0,-1)}$, $\mathcal{R}_{(1,-1)}$, $\mathcal{R}_{(1,0)}$ in this order and return to the starting region $\mathcal{R}_{(1,1)}$. It is also seen that they converge to the same limit cycle passing through the eight regions. As we will see later, it is true that any trajectory $\boldsymbol{\phi}(t, \boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in \mathcal{R}_{(1,1)}$ exhibits the same dynamical behavior as these three. Therefore, we can say that this RNN generates the sequence $S = \{(1, 1), (-1, 1), (-1, -1), (1, -1), (1, 1)\}$. In other words, it generates the sequence $\{(1, 1), (-1, 1), (-1, -1), (1, -1)\}$ repeatedly.

3. Analysis

We present a sufficient condition for the RNN model described by (1)–(3) to generate a given sequence of bipolar vectors satisfying the conditions in Problem 2. We first analyze the dynamical behavior of the RNN model when it consists of only two neurons, and then extend the results to the general case where the number of neurons is not restricted to two.

3.1. Two-neuron case

As the first step, we consider the RNN model consisting of two neurons. The dynamics of the RNN model is described by

$$\begin{cases} \dot{x}_1(t) = g_1(x_1(t), x_2(t)), \\ \dot{x}_2(t) = g_2(x_1(t), x_2(t)), \end{cases} \tag{6}$$

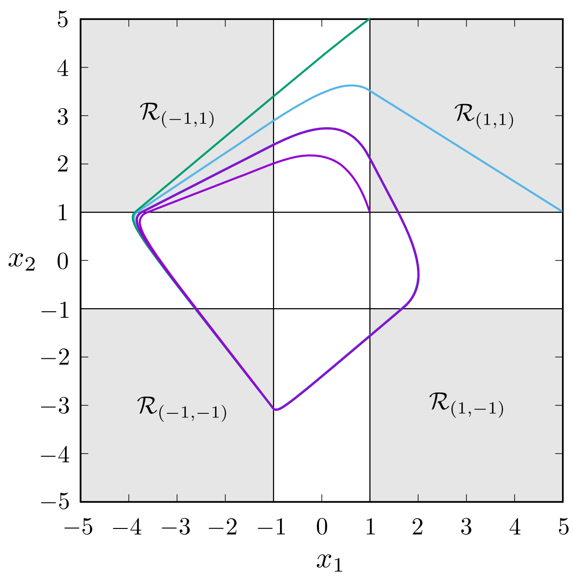


Fig. 1. Three trajectories $\boldsymbol{\phi}(t,(1,1))$, $\boldsymbol{\phi}(t,(5,1))$ and $\boldsymbol{\phi}(t,(1,5))$ generated by the two-neuron RNN with the parameters given by (5).

where

$$g_1(x_1,x_2) = -x_1 + 2f(x_1) + w_{12}f(x_2) + b_1,$$
$$g_2(x_1,x_2) = -x_2 + 2f(x_2) + w_{21}f(x_1) + b_2.$$

The next lemma provides a sufficient condition for the trajectory $\phi(t,(1,1))$ of the state vector of the RNN model to reach $\mathcal{R}_{(-1,1)}$.

Lemma 1. *If the parameters of the RNN model described by (6) satisfy*

$$w_{12} + b_1 < -1, \tag{7}$$
$$w_{21} + b_2 > -1, \tag{8}$$
$$b_2 \geq -1, \tag{9}$$

then the trajectory $\boldsymbol{\phi}(t,(1,1))$ of the state vector passes through the interior of $\mathcal{R}_{(0,1)}$ and reaches $\mathcal{R}_{(-1,1)}$.

Proof. It follows from (7) and (8) that

$$g_1(1,1) = 1 + w_{12} + b_1 < 0,$$
$$g_2(1,1) = 1 + w_{21} + b_2 > 0.$$

Hence the trajectory $\boldsymbol{\phi}(t,(1,1))$ first enters the region $\mathcal{R}_{(0,1)}$. As long as $\boldsymbol{\phi}(t,(1,1)) \in \mathcal{R}_{(0,1)}$, the trajectory $\boldsymbol{\phi}(t,(1,1))$ is explicitly expressed as

$$\phi_1(t,(1,1)) = (w_{12} + b_1 + 1)e^t - w_{12} - b_1, \tag{10}$$

$$\phi_2(t,(1,1)) = \left[\frac{w_{21}}{2}(w_{12} + b_1 - 1) - b_2 - 1\right]e^{-t} + \frac{w_{21}}{2}(w_{12} + b_1 + 1)e^t - w_{21}(w_{12} + b_1) + b_2 + 2. \tag{11}$$

We see from (7) that the right-hand side of (10) is a monotone decreasing function of t, and takes the value of -1 when

$$t = t^* := \ln\left(\frac{w_{12} + b_1 - 1}{w_{12} + b_1 + 1}\right). \tag{12}$$

In order to understand the behavior of $\phi_2(t,(1,1))$, we consider the time derivative of it, which is given by

$$\begin{aligned}\dot{\phi}_2(t,(1,1)) &= -\left[\frac{w_{21}}{2}(w_{12} + b_1 - 1) - b_2 - 1\right]e^{-t} \\ &\quad + \frac{w_{21}}{2}(w_{12} + b_1 + 1)e^t \\ &= \frac{e^{-t}}{2}\left[w_{21}(w_{12} + b_1 + 1)e^{2t}\right. \\ &\quad \left. - w_{21}(w_{12} + b_1 - 1) + 2(b_2 + 1)\right].\end{aligned} \tag{13}$$

Taking the assumptions (7)–(9) into account, we can make the following observations: (i) if $w_{21} < 0$ then the right-hand side of (13) is monotone increasing for $t \geq 0$ because it follows from (13), (7) and (8) that

$$\begin{aligned}&\frac{e^{-t}}{2}\left[w_{21}(w_{12} + b_1 + 1)e^{2t} - w_{21}(w_{12} + b_1 - 1) + 2(b_2 + 1)\right] \\ &\geq \frac{e^{-t}}{2}\left[w_{21}(w_{12} + b_1 + 1) - w_{21}(w_{12} + b_1 - 1) + 2(b_2 + 1)\right] \\ &= e^{-t}(w_{21} + b_2 + 1) \\ &> 0,\end{aligned}$$

(ii) if $w_{21} = 0$ then the right-hand side of (13) is monotone nondecreasing for $t \geq 0$ because it follows from (9) and (13) that

$$\frac{e^{-t}}{2}\left[w_{21}(w_{12} + b_1 + 1)e^{2t} - w_{21}(w_{12} + b_1 - 1)\right. \left. + 2(b_2 + 1)\right] = e^{-t}(b_2 + 1) \geq 0$$

and (iii) if $w_{21} > 0$ then the right-hand side of (13) is monotone increasing for $0 \leq t < T$ where

$$T := \frac{1}{2}\ln\left(\frac{-w_{21}(w_{12} + b_1 - 1) + 2(b_2 + 1)}{-w_{21}(w_{12} + b_1 + 1)}\right)$$

and monotone decreasing for $t > T$. Furthermore, substituting (12) into (11), we have

$$\phi_2(t^*,(1,1)) = 1 - \frac{2(b_2 + 1)}{w_{12} + b_1 - 1}$$

which is greater than or equal to 1 by (7) and (9). From these observations, we can say that $\phi_1(t,(1,1))$ decreases monotonically for $0 \leq t \leq t^*$ and becomes -1 when $t = t^*$, ϕ_2(t,(1,1)) does not decrease for $0 \leq t \leq t^*$ if $w_{21} \leq 0$ or if $w_{21} > 0$ and $T \geq t^*$, and $\phi_2(t,(1,1))$ increases monotonically for $0 \leq t < T$ and then decreases monotonically for $T < t \leq t^*$ if $w_{21} > 0$ and $T < t^*$. The behavior of the trajectory $\phi(t,(1,1))$ in $\mathcal{R}_{(0,1)}$ is shown in Fig. 2. Therefore, if the assumptions (7)–(9) hold then $\phi_2(t,(1,1)) \geq 1$ for all $t \in [0,t^*]$, which means that $\boldsymbol{\phi}(t,(1,1))$ is in the interior of $\mathcal{R}_{(0,1)}$ for all $t \in (0,t^*)$ and $\boldsymbol{\phi}(t^*,(1,1)) \in \mathcal{R}_{(-1,1)}$. This completes the proof. □

A key point in the proof of Lemma 1 is that $\phi_2(t^*,(1,1))$ is expressed in a simple formula. We now show that this occurs only when $w_{11} = 2$. Let us first consider the case where $w_{11} \neq 1$ and w_{22} is not necessarily equal to 2. In this case, $\boldsymbol{\phi}(t,(1,1))$ enters the region $\mathcal{R}_{(0,1)}$ if $w_{11} > 1$, $w_{11} + w_{12} + b_1 - 1 < 0$ and $w_{21} + w_{22} + b_2 - 1 > 0$. Moreover, as long as $\boldsymbol{\phi}(t,(1,1)) \in \mathcal{R}_{(0,1)}$, the trajectory $\boldsymbol{\phi}(t,(1,1))$ is explicitly expressed as

$$\phi_1(t,(1,1)) = \left(\frac{w_{12} + b_1}{w_{11} - 1} + 1\right)e^{(w_{11}-1)t} - \frac{w_{12} + b_1}{w_{11} - 1}, \tag{14}$$

$$\begin{aligned}\phi_2(t,(1,1)) &= \left[\frac{w_{21}}{w_{11}}(w_{12} + b_1 - 1) - (w_{22} + b_2 - 1)\right]e^{-t} \\ &\quad + \frac{w_{21}}{w_{11}(w_{11} - 1)}(w_{11} + w_{12} + b_1 - 1)e^{(w_{11}-1)t} \\ &\quad - \frac{w_{21}}{w_{11} - 1}(w_{12} + b_1) + w_{22} + b_2.\end{aligned} \tag{15}$$

The right-hand side of (14) takes the value of -1 when

$$t = \frac{1}{w_{11} - 1}\ln\left(\frac{-w_{11} + w_{12} + b_1 + 1}{w_{11} + w_{12} + b_1 - 1}\right).$$

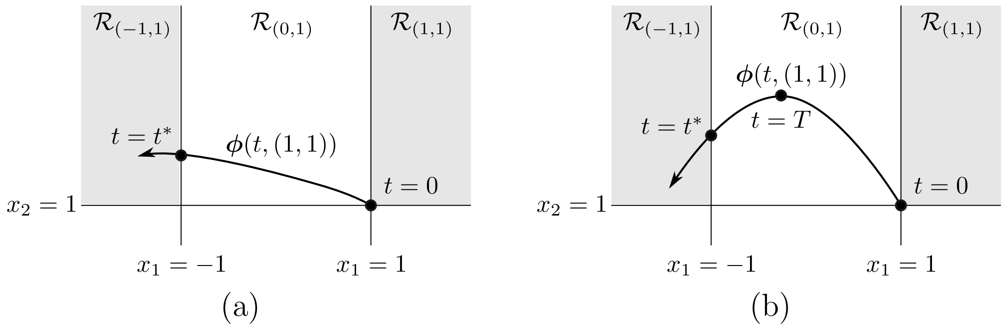


Fig. 2. Behavior of the trajectory $\boldsymbol{\phi}(t,(1,1))$ in $\mathcal{R}_{(0,1)}$. (a) $\phi_2(t,(1,1))$ does not decrease for $0 \le t \le t^*$ if $w_{21} \le 0$ or if $w_{21} > 0$ and $T \ge t^*$. (b) $\phi_2(t,(1,1))$ increases for $0 \le t < T$ and decreases for $T < t \le t^*$ if $w_{21} > 0$ and $T < t^*$.

Substituting this into the right-hand side of (15), we have

$$\left[\frac{w_{21}}{w_{11}}(w_{12}+b_1-1)-(w_{22}+b_2-1)\right] \times \left(\frac{w_{11}+w_{12}+b_1-1}{-w_{11}+w_{12}+b_1+1}\right)^{1/(w_{11}-1)} - \frac{w_{21}}{w_{11}}(w_{12}+b_1+1)+w_{22}+b_2$$

which can be simplified only when $w_{11} = 2$. Let us next consider the case where $w_{11} = 1$ and w_{22} is not necessarily equal to 2. In this case, $\boldsymbol{\phi}(t,(1,1))$ enters the region $\mathcal{R}_{(0,1)}$ if $w_{12}+b_1 < 0$ and $w_{21}+w_{22}+b_2-1 > 0$. As long as $\boldsymbol{\phi}(t,(1,1)) \in \mathcal{R}_{(0,1)}$, the trajectory $\boldsymbol{\phi}(t,(1,1))$ is explicitly expressed as

$$\phi_1(t,(1,1)) = (w_{12}+b_1)t+1, \tag{16}$$

$$\begin{aligned}\phi_2(t,(1,1)) = {} & [w_{21}(w_{12}+b_1-1)-(w_{22}+b_2-1)]\,e^{-t} \\ & + w_{21}(w_{12}+b_1)t - w_{21}(w_{12}+b_1-1)+w_{22}+b_2.\end{aligned} \tag{17}$$

The right-hand side of (16) takes the value of -1 when $t = -2/(w_{12}+b_1)$. Substituting this into the right-hand side of (17), we have

$$[w_{21}(w_{12}+b_1-1)-(w_{22}+b_2-1)]\,e^{2/(w_{12}+b_1)} - w_{21}(w_{12}+b_1+1)+w_{22}+b_2$$

which cannot be simplified further.

Using Lemma 1, we obtain the next lemma which shows that all trajectories $\phi(t,\boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in \mathcal{R}_{(1,1)}$ reach $\mathcal{R}_{(-1,1)}$ under the same condition.

Lemma 2. *If the parameters of the RNN model described by (6) satisfy (7)–(9) then for any $\boldsymbol{x}^0 \in \mathcal{R}_{(1,1)}$ the trajectory $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ of the state vector passes through the interior of $\mathcal{R}_{(0,1)}$ and reaches $\mathcal{R}_{(-1,1)}$.*

Proof. Let $\boldsymbol{x}^0$ be any point in $\mathcal{R}_{(1,1)}$. As long as $\boldsymbol{\phi}(t,\boldsymbol{x}^0) \in \mathcal{R}_{(1,1)}$, $\phi_1(t,\boldsymbol{x}^0)$ decreases monotonically because

$$\begin{aligned} g_1(x_1,x_2)|_{(x_1,x_2)\in\mathcal{R}_{(1,1)}} &= -x_1+2+w_{12}+b_1|_{x_1\ge 1} \\ &\le w_{12}+b_1+1 < 0\end{aligned}$$

follows from (7). Also, $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ does not move from $\mathcal{R}_{(1,1)}$ to $\mathcal{R}_{(1,0)}$ nor $\mathcal{R}_{(0,0)}$ because

$$g_2(x_1,x_2)|_{(x_1,x_2)\in\mathcal{R}_{(1,1)},\,x_2=1} = -1+2+w_{21}+b_2 = w_{21}+b_2+1 > 0$$

follows from (8). Hence $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ moves from $\mathcal{R}_{(1,1)}$ to $\mathcal{R}_{(0,1)}$. As long as $\boldsymbol{\phi}(t,\boldsymbol{x}^0) \in \mathcal{R}_{(0,1)}$, $\phi_1(t,\boldsymbol{x}^0)$ decreases monotonically because

$$\begin{aligned} g_1(x_1,x_2)|_{(x_1,x_2)\in\mathcal{R}_{(0,1)}} &= -x_1+2x_1+w_{12}+b_1|_{|x_1|<1} \\ &\le w_{12}+b_1+1 < 0\end{aligned}$$

follows from (7). In addition, due to the uniqueness of the solution of (6), $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ does not intersect $\boldsymbol{\phi}(t,(1,1))$. Since $\boldsymbol{\phi}(t,(1,1))$ passes through the interior of $\mathcal{R}_{(0,1)}$ and reaches $\mathcal{R}_{(-1,1)}$ as shown in Lemma 1, so does $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$. □

Lemma 2 is generalized as follows.

Lemma 3. *Let $\boldsymbol{\alpha} = (\alpha_1,\alpha_2)$ and $\boldsymbol{\alpha}' = (-\alpha_1,\alpha_2)$ be two bipolar vectors that differ only in the first entry. If the parameters of the RNN model described by (6) satisfy*

$$\alpha_1\alpha_2 w_{12} + \alpha_1 b_1 < -1, \tag{18}$$

$$\alpha_1\alpha_2 w_{21} + \alpha_2 b_2 > -1, \tag{19}$$

$$\alpha_2 b_2 \ge -1, \tag{20}$$

then for any $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}}$ the trajectory $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ of the state vector passes through the interior of $\mathcal{R}_{(\boldsymbol{\alpha}+\boldsymbol{\alpha}')/2}$ and reaches $\mathcal{R}_{\boldsymbol{\alpha}'}$.

Proof. Let us change the variables from $x_1(t)$ and $x_2(t)$ to $\tilde{x}_1(t) = \alpha_1 x_1(t)$ and $\tilde{x}_2(t) = \alpha_2 x_2(t)$. Then we can rewrite (6) as

$$\begin{cases} \dot{\tilde{x}}_1(t) = \tilde{g}_1(\tilde{x}_1(t),\tilde{x}_2(t)), \\ \dot{\tilde{x}}_2(t) = \tilde{g}_2(\tilde{x}_1(t),\tilde{x}_2(t)), \end{cases} \tag{21}$$

where

$$\tilde{g}_1(\tilde{x}_1,\tilde{x}_2) = -\tilde{x}_1 + 2f(\tilde{x}_1) + \alpha_1\alpha_2 w_{12} f(\tilde{x}_2) + \alpha_1 b_1, \tag{22}$$

$$\tilde{g}_2(\tilde{x}_1,\tilde{x}_2) = -\tilde{x}_2 + 2f(\tilde{x}_2) + \alpha_1\alpha_2 w_{21} f(\tilde{x}_1) + \alpha_2 b_2. \tag{23}$$

It follows from Lemma 2 that if the parameters in (22) and (23) satisfy (18)–(20) then for any $\tilde{\boldsymbol{x}}^0 \in \mathcal{R}_{(1,1)}$ the trajectory $\boldsymbol{\phi}(t,\tilde{\boldsymbol{x}}^0)$ of the RNN model (21) passes through the interior of $\mathcal{R}_{(0,1)}$ and reaches $\mathcal{R}_{(-1,1)}$. This means that for any $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}}$ the trajectory $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ of the RNN (6) passes through the interior of $\mathcal{R}_{(\boldsymbol{\alpha}+\boldsymbol{\alpha}')/2}$ and reaches $\mathcal{R}_{\boldsymbol{\alpha}'}$. □

The next lemma immediately follows from Lemma 3.

Lemma 4. *Let $\boldsymbol{\alpha} = (\alpha_1,\alpha_2)$ and $\boldsymbol{\alpha}' = (\alpha_1,-\alpha_2)$ be two bipolar vectors that differ only in the second entry. If the parameters of the RNN model described by (6) satisfy*

$$\alpha_1\alpha_2 w_{12} + \alpha_1 b_1 > -1,$$

$$\alpha_1\alpha_2 w_{21} + \alpha_2 b_2 < -1,$$

$$\alpha_1 b_1 \ge -1,$$

then for any $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}}$ the trajectory $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ of the state vector passes through the interior of $\mathcal{R}_{(\boldsymbol{\alpha}+\boldsymbol{\alpha}')/2}$ and reaches $\mathcal{R}_{\boldsymbol{\alpha}'}$.

Example 1. We find the parameters of the RNN model described by (6) so that it generates the sequence $S = \{(1,1),(-1,1),$

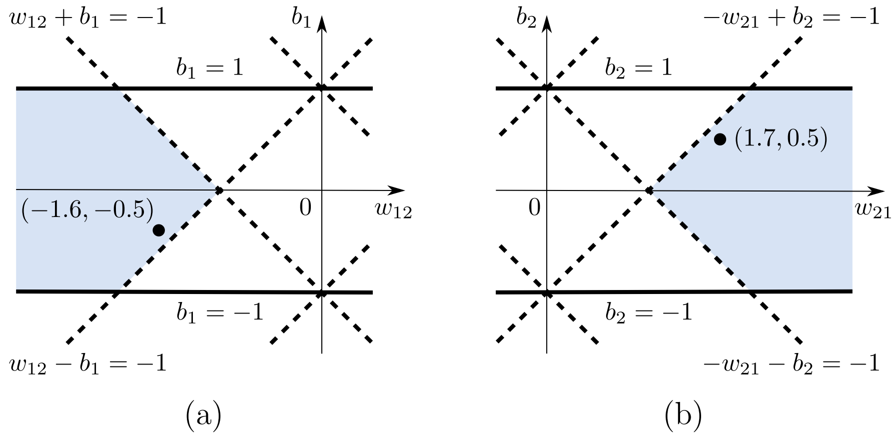


Fig. 3. Parameter regions corresponding to the conditions in Example 1 and the points corresponding to the parameters given in (5) in (a) $w_{12}-b_1$ plane and (b) $w_{21}-b_2$ plane.

$(-1,-1),(1,-1),(1,1)\}$. It follows from Lemmas 3 and 4 that if the parameters of the RNN model satisfy

$$\begin{array}{rrr} w_{12}+b_1<-1, & w_{21}+b_2>-1, & b_2\geq -1, \\ -w_{12}-b_1>-1, & -w_{21}+b_2<-1, & -b_1\geq -1, \\ w_{12}-b_1<-1, & w_{21}-b_2>-1, & -b_2\geq -1, \\ -w_{12}+b_1>-1, & -w_{21}-b_2<-1, & b_1\geq -1, \end{array}$$

then for any $\boldsymbol{x}^0 \in \mathcal{R}_{(1,1)}$ the trajectory $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ of the state vector visits $\mathcal{R}_{(-1,1)}, \mathcal{R}_{(-1,-1)}, \mathcal{R}_{(1,-1)}, \mathcal{R}_{(1,1)}$ in this order. Since the parameters $w_{12}=-1.6$, $w_{21}=1.7$, $b_1=-0.5$ and $b_2=0.5$ given in (5) satisfy all of these inequalities, it is theoretically guaranteed that the RNN with these parameters generates the sequence S. The regions of the parameters corresponding to the above inequalities are the points corresponding to the parameters given in (5) are shown in Fig. 3.

3.2. General case

We extend the results obtained in the previous subsection to the general case where n is not restricted to two.

Lemma 5. *Let* $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\ldots,\alpha_n)$ *and* $\boldsymbol{\alpha}'=(\alpha_1,\ldots,\alpha_{i^*-1},-\alpha_{i^*},\alpha_{i^*+1},\ldots,\alpha_n)$ *be two bipolar vectors that differ only in the* i^**-th entry. If the parameters of the RNN model described by (1)–(3) satisfy*

$$\alpha_{i^*}\left(\sum_{j=1,j\neq i^*}^{n}\alpha_j w_{i^*j}+b_{i^*}\right)<-1, \tag{24}$$

$$\alpha_i\left(\sum_{j=1,j\neq i}^{n}\alpha_j w_{ij}+b_i\right)>-1,\ \forall i\neq i^*, \tag{25}$$

$$\alpha_i\left(\sum_{j=1,j\neq i,i^*}^{n}\alpha_j w_{ij}+b_i\right)\geq -1,\ \forall i\neq i^*, \tag{26}$$

then for any $\boldsymbol{x}^0\in\mathcal{R}_{\boldsymbol{\alpha}}$ *the trajectory* $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ *of the state vector passes through the interior of* $\mathcal{R}_{(\boldsymbol{\alpha}+\boldsymbol{\alpha}')/2}$ *and reaches* $\mathcal{R}_{\boldsymbol{\alpha}'}$.

Proof. The dynamics of the RNN model is expressed as

$$\dot{x}_i(t)=g_i(x_1(t),x_2(t),\ldots,x_n(t)),\ i=1,2,\ldots,n$$

where

$$g_i(x_1,x_2,\ldots,x_n)=-x_i+2f(x_i)+\sum_{j=1,j\neq i}^{n}w_{ij}f(x_j)+b_i,\ i=1,2,\ldots,n.$$

In what follows, we assume without loss of generality that $i^*=1$. As long as $\boldsymbol{\phi}(t,\boldsymbol{x}^0)\in\mathcal{R}_{\boldsymbol{\alpha}}$, $\alpha_1\phi_1(t,\boldsymbol{x}^0)$ decreases monotonically because

$$\begin{aligned} \alpha_1 g_1(x_1,x_2,\ldots,x_n)|_{\boldsymbol{x}\in\mathcal{R}_{\boldsymbol{\alpha}}} &= \alpha_1\left(-x_1+2\alpha_1+\sum_{j=2}^{n}w_{1j}\alpha_j+b_1\right)\Bigg|_{\alpha_1x_1\geq 1} \\ &\leq 1+\alpha_1\left(\sum_{j=2}^{n}w_{1j}\alpha_j+b_1\right) \\ &<0 \end{aligned}$$

follows from (24). Also, as long as $\boldsymbol{\phi}(t,\boldsymbol{x}^0)\in\mathcal{R}_{\boldsymbol{\alpha}}$, $\alpha_i\phi_i(t,\boldsymbol{x}^0)$ is not less than 1 for $i=2,3,\ldots,n$ because

$$\begin{aligned} \alpha_i g_i(x_1,x_2,\ldots,x_n)|_{\boldsymbol{x}\in\mathcal{R}_{\boldsymbol{\alpha}},x_i=\alpha_i} &= \alpha_i\left(-\alpha_i+2\alpha_i+\sum_{j=1,j\neq i}^{n}w_{ij}\alpha_j+b_i\right) \\ &= 1+\alpha_i\left(\sum_{j=1,j\neq i}^{n}w_{ij}\alpha_j+b_i\right) \\ &>0 \end{aligned}$$

follows from (25). Hence $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ moves from $\mathcal{R}_{\boldsymbol{\alpha}}$ to $\mathcal{R}_{(\boldsymbol{\alpha}+\boldsymbol{\alpha}')/2}$. As long as $\boldsymbol{\phi}(t,\boldsymbol{x}^0)\in\mathcal{R}_{(\boldsymbol{\alpha}+\boldsymbol{\alpha}')/2}$, $\alpha_1\phi_1(t,\boldsymbol{x}^0)$ decreases monotonically because

$$\begin{aligned} &\alpha_1 g_1(x_1,x_2,\ldots,x_n)|_{\boldsymbol{x}\in\mathcal{R}_{(\boldsymbol{\alpha}+\boldsymbol{\alpha}')/2}} \\ &=\alpha_1\left(-x_1+2x_1+\sum_{j=2}^{n}w_{1j}\alpha_j+b_1\right)\Bigg|_{|x_1|<1} \\ &<1+\alpha_1\left(\sum_{j=2}^{n}w_{1j}\alpha_j+b_1\right) \\ &<0 \end{aligned}$$

follows from (24). Now we shall prove that $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ moves from $\mathcal{R}_{(\boldsymbol{\alpha}+\boldsymbol{\alpha}')/2}$ to $\mathcal{R}_{\boldsymbol{\alpha}'}$ by contradiction. Assume this is not true. Then there exist a positive number t^* and an integer $i'\in\{2,3,\ldots,n\}$ such that $\boldsymbol{\phi}(t,\boldsymbol{x}^0)\in\mathcal{R}_{\boldsymbol{\alpha}}\cup\mathcal{R}_{(\boldsymbol{\alpha}+\boldsymbol{\alpha}')/2}$ for all $t\in[0,t^*]$, $-1<\alpha_1\phi_1(t^*,\boldsymbol{x}^0)<1$ and $\alpha_{i'}\phi_{i'}(t^*,\boldsymbol{x}^0)=1$. We assume without loss of generality that $i'=2$. Then the behavior of neurons 1 and 2 for $0\leq t\leq t^*$ is described by

$$\begin{cases} \dot{x}_1(t)=-x_1(t)+2f(x_1(t))+w_{12}f(x_2(t))+\tilde{b}_1, \\ \dot{x}_2(t)=-x_2(t)+2f(x_2(t))+w_{21}f(x_1(t))+\tilde{b}_2 \end{cases} \tag{27}$$

where $\tilde{b}_1 = \sum_{j=3}^{n} w_{1j}\alpha_j + b_1$ and $\tilde{b}_2 = \sum_{j=3}^{n} w_{2j}\alpha_j + b_2$. We see from (24)–(26) that the parameters of (27) satisfy the following inequalities:

$$\alpha_1\alpha_2 w_{12} + \alpha_1\tilde{b}_1 < -1,$$
$$\alpha_1\alpha_2 w_{21} + \alpha_2\tilde{b}_2 > -1,$$
$$\alpha_2\tilde{b}_2 \geq -1.$$

Also, $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}}$ implies that $(x_1^0, x_2^0) \in \mathcal{R}_{(\alpha_1,\alpha_2)}$. From these observations and Lemma 3, we see that $(\phi_1(t,\boldsymbol{x}^0), \phi_2(t,\boldsymbol{x}^0))$ passes through the interior of $\mathcal{R}_{(0,\alpha_2)}$ and reaches $\mathcal{R}_{(-\alpha_1,\alpha_2)}$. However, this contradicts the fact that $-1 < \phi_1(t^*,\boldsymbol{x}^0) < 1$ and $\phi_2(t^*,\boldsymbol{x}^0) = \alpha_2$. Therefore, $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ moves from $\mathcal{R}_{(\boldsymbol{\alpha}+\boldsymbol{\alpha}')/2}$ to $\mathcal{R}_{\boldsymbol{\alpha}'}$. □

The next theorem immediately follows from Lemma 5.

Theorem 1. *Let $S = \{\boldsymbol{\alpha}^0, \boldsymbol{\alpha}^1, \ldots, \boldsymbol{\alpha}^m\}$ be a sequence of n-dimensional bipolar vectors satisfying the following conditions.*

1. *$\boldsymbol{\alpha}^{k+1}$ differs from $\boldsymbol{\alpha}^k$ only in the i_k-th entry for $k = 0, 1, \ldots, m-1$.*
2. *$\boldsymbol{\alpha}^k = \boldsymbol{\alpha}^{k'}$ holds only if $k = 0$ and $k' = m$.*

If the parameters of the RNN model described by (1)–(3) satisfy

$$\alpha_{i_k}^k \left(\sum_{j=1, j\neq i_k}^{n} \alpha_j^k w_{i_k j} + b_{i_k} \right) < -1 \tag{28}$$

$$\alpha_i^k \left(\sum_{j=1, j\neq i}^{n} \alpha_j^k w_{ij} + b_i \right) > -1, \ \forall i \neq i_k \tag{29}$$

$$\alpha_i^k \left(\sum_{j=1, j\neq i_k, i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1, \ \forall i \neq i_k \tag{30}$$

for $k = 0, 1, \ldots, m-1$ then for any $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}^0}$ the trajectory $\boldsymbol{\phi}(t,\boldsymbol{x}^0)$ of the state vector passes through $\mathcal{R}_{\boldsymbol{\alpha}^0}, \mathcal{R}_{(\boldsymbol{\alpha}^0+\boldsymbol{\alpha}^1)/2}, \mathcal{R}_{\boldsymbol{\alpha}^1}, \mathcal{R}_{(\boldsymbol{\alpha}^1+\boldsymbol{\alpha}^2)/2}, \mathcal{R}_{\boldsymbol{\alpha}^2}, \ldots, \mathcal{R}_{(\boldsymbol{\alpha}^{m-1}+\boldsymbol{\alpha}^m)/2}$ and reaches $\mathcal{R}_{\boldsymbol{\alpha}^m}$.

The sufficient condition given in Theorem 1 is expressed as a system of linear inequalities in the parameters of the RNN model, and thus easy to test whether it has a solution or not. We should note that this simple form of the condition is due to the assumption that $w_{ii} = 2$ for $i = 1, 2, \ldots, n$. A more general condition may be obtained if this assumption is removed, but it is not clear whether the condition can be expressed in a simple form.

4. Design

We consider the problem of finding the parameters of the RNN model described by (1)–(3) generating a given sequence $S = \{\boldsymbol{\alpha}^0, \boldsymbol{\alpha}^1, \ldots, \boldsymbol{\alpha}^m\}$ of bipolar vectors satisfying the conditions given in Theorem 1. A simple way is to solve the system of linear inequalities described by (28)–(30). However, when the system of inequalities is feasible, it has an infinite number of solutions in general. It is thus natural to formulate the problem as a constrained optimization problem in which an objective function has to be minimized or maximized subject to the linear inequality constraints (28)–(30). In this paper, we consider three types of constrained optimization problems. One is formulated as a convex quadratic programming (QP) problem and the others are linear programming (LP) problems. We examine the advantages and disadvantages of these three approaches by using an example of sequence of bipolar vectors.

4.1. Minimization of sum of squares of parameters

The first approach is to minimize the sum of squares of the parameters of the RNN model described by (1)–(3) under the constraints (28)–(30). By doing so, we may be able to make the values of unnecessary weights and biases zero, which means that we obtain an RNN with a simple structure. This approach is formulated as the optimization problem:

$$\begin{array}{ll}\text{minimize} & \sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} w_{ij}^2 + \sum_{i=1}^{n} b_i^2 \\ \text{subject to} & \alpha_{i_k}^k \left(\sum_{j=1, j\neq i_k}^{n} \alpha_j^k w_{i_k j} + b_{i_k} \right) \leq -1-\epsilon, \ \forall k, \\ & \alpha_i^k \left(\sum_{j=1, j\neq i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1+\epsilon, \ \forall i \neq i_k, \ \forall k, \\ & \alpha_i^k \left(\sum_{j=1, j\neq i_k, i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1, \ \forall i \neq i_k, \ \forall k,\end{array} \tag{31}$$

where ϵ is a small positive constant.

Let $\mathcal{K}_i(S) := \{k \mid i_k = i\}$ and $\bar{\mathcal{K}}_i(S) := \{k \mid i_k \neq i\}$. It is clear that $\mathcal{K}_i(S) \cap \bar{\mathcal{K}}_i(S) = \emptyset$ and $\mathcal{K}_i(S) \cup \bar{\mathcal{K}}_i(S) = \{0, 1, \ldots, m-1\}$. Hence the problem (31) can be decomposed into n independent problems of the form:

$$\begin{array}{ll}\text{minimize} & \sum_{j=1, j\neq i}^{n} w_{ij}^2 + b_i^2 \\ \text{subject to} & \alpha_i^k \left(\sum_{j=1, j\neq i_k}^{n} \alpha_j^k w_{ij} + b_i \right) \leq -1-\epsilon, \ \forall k \in \mathcal{K}_i(S), \\ & \alpha_i^k \left(\sum_{j=1, j\neq i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1+\epsilon, \ \forall k \in \bar{\mathcal{K}}_i(S), \\ & \alpha_i^k \left(\sum_{j=1, j\neq i_k, i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1, \ \forall k \in \bar{\mathcal{K}}_i(S).\end{array} \tag{32}$$

Since the objective function is quadratic and convex, and the constraints are linear, the problem (32) is a convex QP problem, which can be easily solved by using a QP solver. For example, an infeasible-interior-point algorithm (Potra, 1996) can find an approximate solution of (32) in polynomial time if it is feasible. In order to derive the time complexity of the algorithm, we convert (32) into the standard form. We first write each variable, which is not restricted to be nonnegative, as the difference of two new nonnegative variables. We then convert each inequality constraint to an equality constraint by adding or subtracting a nonnegative slack variable. Since the number of inequality constraints in (32) is at most $2m$, the number of slack variables needed is at most $2m$. As a result, we obtain the standard form with at most $2(n+m)$ variables and at most $2m$ equality constraints. Therefore, the time complexity of the infeasible-interior-point algorithm is $O((n+m)^4 L)$, where L is the length of a binary string encoding the standard form.

4.2. Minimization of sum of absolute values of parameters

The second approach is to minimize the sum of absolute values of the parameters of the RNN model described by (1)–(3) under the constraints (28)–(30). By replacing the sum of squares in the first approach with the sum of absolute values, we may be able to make the weight matrix and the bias vector sparser in the sense that more entries equal to zero, just like the ℓ_1-norm minimization for recovering sparse signals (see, for example, a survey paper (Zhang et al., 2015) and references therein). This approach is formulated as the optimization problem:

$$\begin{array}{ll}\text{minimize} & \sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} |w_{ij}| + \sum_{i=1}^{n} |b_i| \\ \text{subject to} & \alpha_{i_k}^k \left(\sum_{j=1, j\neq i_k}^{n} \alpha_j^k w_{i_k j} + b_{i_k} \right) \leq -1-\epsilon, \ \forall k, \\ & \alpha_i^k \left(\sum_{j=1, j\neq i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1+\epsilon, \ \forall i \neq i_k, \ \forall k, \\ & \alpha_i^k \left(\sum_{j=1, j\neq i_k, i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1, \ \forall i \neq i_k, \ \forall k,\end{array} \tag{33}$$

where ϵ is a small positive constant. As in the previous approach, the problem (33) can be decomposed into n independent problems of the form:

$$\begin{array}{ll} \text{minimize} & \sum_{j=1,\, j\neq i}^{n} |w_{ij}| + |b_i| \\ \text{subject to} & \alpha_i^k \left(\sum_{j=1, j\neq i_k}^{n} \alpha_j^k w_{ij} + b_i \right) \leq -1-\epsilon,\ \forall k \in \mathcal{K}_i(S), \\ & \alpha_i^k \left(\sum_{j=1, j\neq i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1+\epsilon,\ \forall k \in \bar{\mathcal{K}}_i(S), \\ & \alpha_i^k \left(\sum_{j=1, j\neq i_k, i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1,\ \forall k \in \bar{\mathcal{K}}_i(S). \end{array} \tag{34}$$

Furthermore, this problem can be recast as the optimization problem:

$$\begin{array}{ll} \text{minimize} & \sum_{j=1,\, j\neq i}^{n} (w_{ij}^+ + w_{ij}^-) + b_i^+ + b_i^- \\ \text{subject to} & \alpha_i^k \left(\sum_{j=1, j\neq i_k}^{n} \alpha_j^k (w_{ij}^+ - w_{ij}^-) + b_i^+ - b_i^- \right) \\ & \quad \leq -1-\epsilon,\ \forall k \in \mathcal{K}_i(S), \\ & \alpha_i^k \left(\sum_{j=1, j\neq i}^{n} \alpha_j^k (w_{ij}^+ - w_{ij}^-) + b_i^+ - b_i^- \right) \\ & \quad \geq -1+\epsilon,\ \forall k \in \bar{\mathcal{K}}_i(S), \\ & \alpha_i^k \left(\sum_{j=1, j\neq i_k, i}^{n} \alpha_j^k (w_{ij}^+ - w_{ij}^-) + b_i^+ - b_i^- \right) \\ & \quad \geq -1,\ \forall k \in \bar{\mathcal{K}}_i(S), \\ & w_{ij}^+, w_{ij}^- \geq 0,\ \forall j \neq i, \\ & b_i^+, b_i^- \geq 0, \end{array} \tag{35}$$

where w_{ij}^+, w_{ij}^- $(j = 1, 2, \ldots, i-1, i+1, i+2, \ldots, n)$, b_i^+ and b_i^- are new variables. Since the objective function and the constraints are all linear, the problem (35) is an LP problem, which can be easily solved by using an LP solver. For example, an infeasible-interior-point algorithm (Mizuno, 1994) can find an approximate solution of (35) in polynomial time if it is feasible. Since the standard form of (35) has at most $2(n+m)$ variables and at most $2m$ equality constraints, the time complexity of the algorithm is $O((n+m)^4 L)$, where L is the length of a binary string encoding the standard form.

4.3. Maximization of margin in inequality constraints

The third approach is to maximize the margin in the inequality constraints in order to ensure robustness against parameter variations. This is particularly important for the analog circuit implementation of the RNN because the values of circuit elements differ from the desired ones in general and may change depending on the environment. This approach is formulated as the optimization problem:

$$\begin{array}{ll} \text{maximize} & \sum_{i=1}^{n} \delta_i \\ \text{subject to} & \alpha_{i_k}^k \left(\sum_{j=1, j\neq i_k}^{n} \alpha_j^k w_{i_k j} + b_{i_k} \right) \leq -1-\epsilon-\delta_i,\ \forall k, \\ & \alpha_i^k \left(\sum_{j=1, j\neq i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1+\epsilon+\delta_i,\ \forall i \neq i_k,\ \forall k, \\ & \alpha_i^k \left(\sum_{j=1, j\neq i_k, i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1+\delta_i,\ \forall i \neq i_k,\ \forall k, \\ & |w_{ij}| \leq U,\ \forall i \neq j, \\ & |b_i| \leq U,\ \forall i, \\ & \delta_i \geq 0,\ \forall i, \end{array} \tag{36}$$

where $\delta_1, \delta_2, \ldots, \delta_n$ are additional variables, and ϵ and U are positive constants. Note that the fourth and fifth constraints are necessary because otherwise the problem has no optimal solution. To be more specific, the value of the objective function is not bounded above when the first three constraints are feasible. Note also that the value of ϵ should be sufficiently smaller than that of U.

As in the first and second approaches, the problem (36) can be decomposed into n independent problems of the form:

$$\begin{array}{ll} \text{maximize} & \delta_i \\ \text{subject to} & \alpha_i^k \left(\sum_{j=1, j\neq i_k}^{n} \alpha_j^k w_{ij} + b_i \right) \leq -1-\epsilon-\delta_i,\ \forall k \in \mathcal{K}_i(S), \\ & \alpha_i^k \left(\sum_{j=1, j\neq i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1+\epsilon+\delta_i,\ \forall k \in \bar{\mathcal{K}}_i(S), \\ & \alpha_i^k \left(\sum_{j=1, j\neq i_k, i}^{n} \alpha_j^k w_{ij} + b_i \right) \geq -1+\delta_i,\ \forall k \in \bar{\mathcal{K}}_i(S), \\ & |w_{ij}| \leq U,\ \forall j \neq i, \\ & |b_i| \leq U, \\ & \delta_i \geq 0. \end{array} \tag{37}$$

Since the objective function and the constraints are all linear, the problem (37) is an LP problem, which can be easily solved by using an LP solver. Since the standard form of (37) has at most $4n + 2m + 1$ variables and $2(n + m)$ equality constraints, the infeasible-interior-point algorithm (Mizuno, 1994) can find an approximate solution of (37) in $O((n+m)^4 L)$ time if it is feasible, where L is the length of a binary string encoding the standard form.

4.4. Examples

In order to demonstrate the differences among the proposed three design methods, we apply them to the same sequence of bipolar vectors to obtain three different RNNs, and compare the trajectories of the state vector generated by these RNNs for the same initial condition.

Example 2. Let $S = \{\boldsymbol{\alpha}^0, \boldsymbol{\alpha}^1, \ldots, \boldsymbol{\alpha}^8\}$ with $\boldsymbol{\alpha}^0 = (1, 1, 1, 1)$, $\boldsymbol{\alpha}^1 = (-1, 1, 1, 1)$, $\boldsymbol{\alpha}^2 = (-1, -1, 1, 1)$, $\boldsymbol{\alpha}^3 = (1, -1, 1, 1)$, $\boldsymbol{\alpha}^4 = (1, -1, -1, 1)$, $\boldsymbol{\alpha}^5 = (1, 1, -1, 1)$, $\boldsymbol{\alpha}^6 = (1, 1, -1, -1)$, $\boldsymbol{\alpha}^7 = (1, 1, 1, -1)$ and $\boldsymbol{\alpha}^8 = (1, 1, 1, 1)$. First, using the method to minimize the sum of squares of the parameters, that is, solving the QP problem (32) with $\epsilon = 0.01$ for $i = 1, 2, 3, 4$, we obtain

$$\boldsymbol{W} = \begin{bmatrix} 2.000 & -1.010 & -0.010 & -0.010 \\ 0.505 & 2.000 & -0.505 & 0.000 \\ -0.010 & 0.505 & 2.000 & -0.505 \\ 0.005 & -0.010 & 1.010 & 2.000 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 0.020 \\ 0.000 \\ 0.010 \\ 0.005 \end{bmatrix}. \tag{38}$$

Note that the values of w_{24} and b_2 are zero. The time evolution of $\boldsymbol{\phi}(t, (2, 2, 2, 2))$ for the designed RNN is shown in Fig. 4. Gray boxes indicate the time intervals in which $\boldsymbol{f}(\boldsymbol{\phi}(t, (2, 2, 2, 2)))$ is a bipolar vector contained in the sequence S. It is easily seen that the given sequence S is certainly generated by the RNN. It is also seen that the trajectory does not stay the regions $\mathcal{R}_{\alpha^3}$, $\mathcal{R}_{\alpha^5}$ and $\mathcal{R}_{\alpha^8}$ even for a moment, so these bipolar vectors could go unnoticed. This is because some of the inequality constraints in (31) holds with the equal sign.

Next, using the method to minimize the sum of absolute values of the parameters, that is, solving the LP problem (34) with $\epsilon = 0.01$ for $i = 1, 2, 3, 4$, we obtain

$$\boldsymbol{W} = \begin{bmatrix} 2.000 & -1.010 & -0.010 & -0.010 \\ 1.000 & 2.000 & -0.010 & 0.000 \\ -0.010 & 1.000 & 2.000 & -0.010 \\ 0.010 & -0.010 & 1.010 & 2.000 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 0.020 \\ 0.000 \\ 0.010 \\ 0.000 \end{bmatrix}. \tag{39}$$

Note that the values of w_{24}, b_2 and b_4 are zero. As expected, we obtained a sparser solution than the previous method. The time evolution of $\boldsymbol{\phi}(t, (2, 2, 2, 2))$ for the designed RNN is shown in Fig. 5. It is seen that the given sequence S of bipolar vectors is

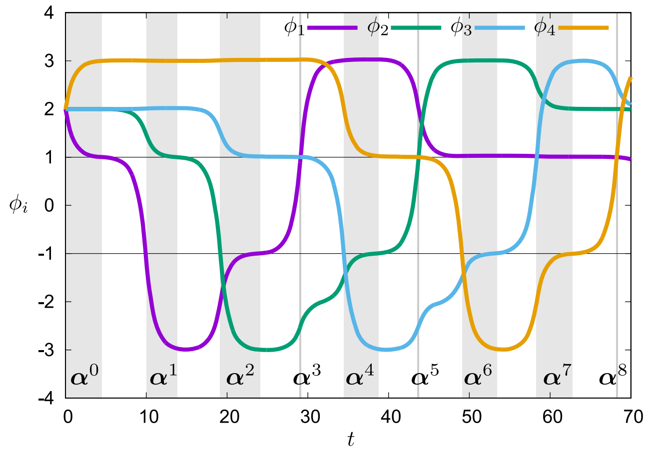


Fig. 4. Time evolution of the entries of $\boldsymbol{\phi}(t,(2,2,2,2))$ generated by the RNN with the parameters given by (38).

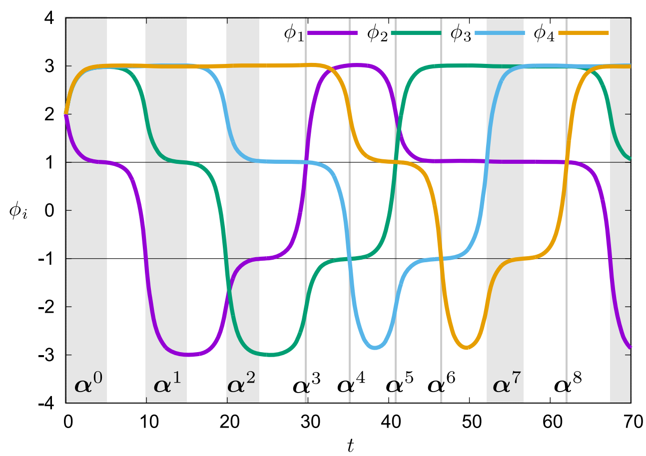


Fig. 5. Time evolution of the entries of $\boldsymbol{\phi}(t,(2,2,2,2))$ generated by the RNN with the parameters given by (39).

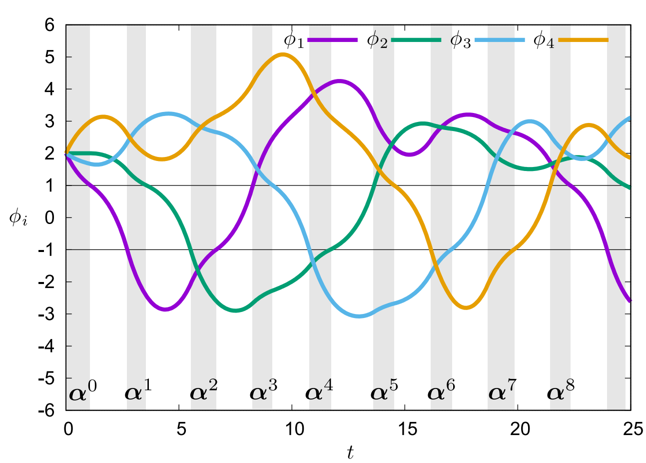


Fig. 6. Time evolution of the entries of $\boldsymbol{\phi}(t,(2,2,2,2))$ generated by the RNN with the parameters given by (40).

certainly generated by the RNN. It is also seen that the trajectory does not stay the regions $\mathcal{R}_{\boldsymbol{\alpha}^3}$, $\mathcal{R}_{\boldsymbol{\alpha}^4}$, $\mathcal{R}_{\boldsymbol{\alpha}^5}$, $\mathcal{R}_{\boldsymbol{\alpha}^6}$ and $\mathcal{R}_{\boldsymbol{\alpha}^8}$ even for a moment, so these bipolar vectors could go unnoticed. The reason for this is the same as in the previous method.

Finally, using the method to maximize the margin in the inequality constraints, that is, solving the LP problem (37) with $\epsilon = 0.01$ and $U = 1.5$ for $i = 1, 2, 3, 4$, we obtain

$$\boldsymbol{W} = \begin{bmatrix} 2.000 & -1.500 & -0.500 & -0.990 \\ 0.670 & 2.000 & -0.670 & 0.330 \\ -1.000 & 0.505 & 2.000 & -1.000 \\ 1.000 & -0.990 & 1.500 & 2.000 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1.490 \\ -0.330 \\ 1.000 \\ -0.010 \end{bmatrix}, \tag{40}$$

and $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_3, \delta_4) = (0.490, 0.330, 0.495, 0.490)$. Note that all entries of $\boldsymbol{W}$ and $\boldsymbol{b}$ take nonzero values. The time evolution of $\boldsymbol{\phi}(t,(2,2,2,2))$ for the designed RNN is shown in Fig. 6. It is seen that the given sequence S of bipolar vectors is certainly generated by the RNN. It is also seen that, unlike the previous two methods, every bipolar vector in S appears for at least a certain amount of time. This is because the first three inequality constraints in (36) are satisfied with the maximum margin. As for the robustness against parameter variations, it can be said from the value of $\boldsymbol{\delta}$ that the RNN with the parameter values changed by Δw_{ij} and Δb_i still generates the sequence S if the following inequalities hold:

$$\max\{|\Delta w_{12}|, |\Delta w_{13}|, |\Delta w_{14}|, |\Delta b_1|\} \le 0.490/4 = 0.1225,$$
$$\max\{|\Delta w_{21}|, |\Delta w_{23}|, |\Delta w_{24}|, |\Delta b_2|\} \le 0.330/4 = 0.0825,$$
$$\max\{|\Delta w_{31}|, |\Delta w_{32}|, |\Delta w_{34}|, |\Delta b_3|\} \le 0.495/4 = 0.12375,$$
$$\max\{|\Delta w_{41}|, |\Delta w_{42}|, |\Delta w_{43}|, |\Delta b_4|\} \le 0.490/4 = 0.1225.$$

We give another example to show that the proposed design methods can be easily extended to the case where a single RNN generates different sequences of bipolar vectors depending on the initial output vector.

Example 3. Let us consider three sequences of bipolar vectors $S^1 = \{\boldsymbol{\alpha}^{1,0}, \boldsymbol{\alpha}^{1,1}, \ldots, \boldsymbol{\alpha}^{1,4}\}$ with $\boldsymbol{\alpha}^{1,0} = (1,1,1)$, $\boldsymbol{\alpha}^{1,1} = (-1,1,1)$, $\boldsymbol{\alpha}^{1,2} = (-1,-1,1)$, $\boldsymbol{\alpha}^{1,3} = (1,-1,1)$ and $\boldsymbol{\alpha}^{1,4} = (1,1,1) = \boldsymbol{\alpha}^{1,0}$, $S^2 = \{\boldsymbol{\alpha}^{2,0}, \boldsymbol{\alpha}^{2,1}\}$ with $\boldsymbol{\alpha}^{2,0} = (-1,1,-1)$ and $\boldsymbol{\alpha}^{2,1} = (1,1,-1)$, and $S^3 = \{\boldsymbol{\alpha}^{3,0}, \boldsymbol{\alpha}^{3,1}, \boldsymbol{\alpha}^{3,2}\}$ with $\boldsymbol{\alpha}^{3,0} = (-1,-1,-1)$, $\boldsymbol{\alpha}^{3,1} = (1,-1,-1)$ and $\boldsymbol{\alpha}^{3,2} = (1,1,-1) = \boldsymbol{\alpha}^{2,1}$. For each $l \in \{1,2,3\}$ and $k \in \{0,1,\ldots,m_l-2\}$ where m_l is the length of S^l, $\boldsymbol{\alpha}^{l,k+1}$ differs from $\boldsymbol{\alpha}^{l,k}$ in one and only one entry. Let the index of the entry be denoted by $i_k(S^l)$. Also, let $\mathcal{K}_i(S^l) = \{k, \mid i_k(S^l) = i\}$ and $\bar{\mathcal{K}}_i(S^l) = \{k \mid i_k(S^l) \ne i\}$. Then the maximization of the margins in the inequality constraints is formulated as n independent optimization problems of the form:

$$\begin{aligned}
\text{maximize} \quad & \delta_i \\
\text{subject to} \quad & \alpha_i^{l,k}\left(\textstyle\sum_{j=1, j\ne i_k(S^l)}^{n} \alpha_j^{l,k} w_{ij} + b_i\right) \\
& \quad \le -1 - \epsilon - \delta_i, \ \forall k \in \mathcal{K}_i(S^l), \ l = 1,2,3, \\
& \alpha_i^{l,k}\left(\textstyle\sum_{j=1, j\ne i}^{n} \alpha_j^{l,k} w_{ij} + b_i\right) \\
& \quad \ge -1 + \epsilon + \delta_i, \ \forall k \in \bar{\mathcal{K}}_i(S^l), \ l = 1,2,3, \\
& \alpha_i^{l,k}\left(\textstyle\sum_{j=1, j\ne i_k(S^l), i}^{n} \alpha_j^{l,k} w_{ij} + b_i\right) \\
& \quad \ge -1 + \delta_i, \ \forall k \in \bar{\mathcal{K}}_i(S^l), \ l = 1,2,3, \\
& |w_{ij}| \le U, \ \forall j \ne i, \\
& |b_i| \le U, \\
& \delta_i \ge 0.
\end{aligned}$$

Solving this problem with $\epsilon = 0.01$ and $U = 1.5$ for $i = 1, 2, 3$, we obtain

$$\boldsymbol{W} = \begin{bmatrix} 2.000 & -1.500 & -1.500 \\ 1.33\dot{6} & 2.000 & -0.33\dot{6} \\ 0.000 & 0.000 & 2.000 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1.500 \\ 0.33\dot{6} \\ 0.000 \end{bmatrix}, \tag{41}$$

and $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_3) = (0.490, 0.32\dot{6}, 0.990)$, where $\dot{6}$ means that the digit 6 repeats infinitely. The time evolution of $\boldsymbol{\phi}(t,(3,3,3))$,

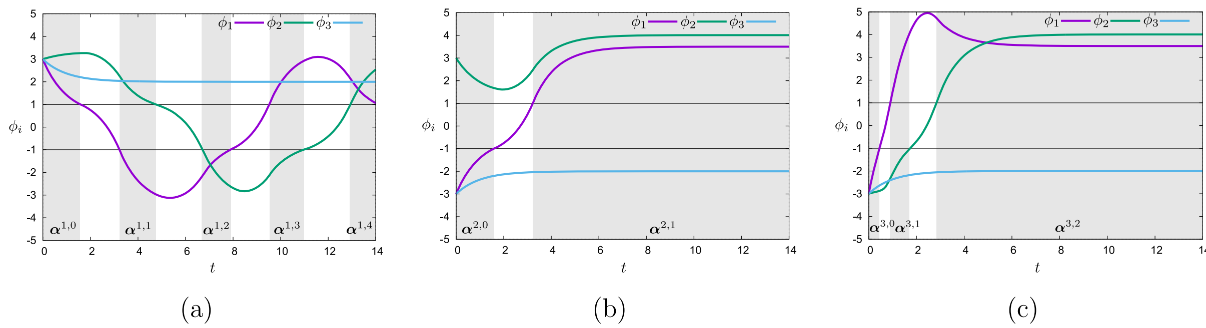


Fig. 7. Time evolution of the entries of (a) $\boldsymbol{\phi}(t,(3,3,3))$, (b) $\boldsymbol{\phi}(t,(-3,3,-3))$ and (c) $\boldsymbol{\phi}(t,(-3,-3,-3))$ generated by the RNN with the parameters given by (41).

$\boldsymbol{\phi}(t,(-3,3,-3))$ and $\boldsymbol{\phi}(t,(-3,-3,-3))$ for the designed RNN is shown in Fig. 7. It is seen that the given sequences S^1, S^2 and S^3 of bipolar vectors are certainly generated by the designed RNN. It is also seen that $\boldsymbol{\phi}(t,(-3,-3,3))$ and $\boldsymbol{\phi}(t,(-3,-3,-3))$ converge to the same constant point $(3.5, 4.01, -2)$ which is a stable equilibrium point in $\mathcal{R}_{(1,1,-1)}$.

5. Some classes of realizable sequences of bipolar vectors

Theorem 1 provides a sufficient condition for a given sequence of bipolar vectors to be generated by the RNN model described by (1)–(3). However, it is not clear what kind of sequence of bipolar vectors guarantees the existence of a solution to the system of inequalities (28)–(30). In this section, we give two types of such sequences.

The sequence of the first type starts with the vector of all 1's. Next the first entry changes from 1 to -1. Then -1 spreads from the first entry to the last entry as k increases, and the vector of all -1's is obtained when $k=n$. In the next vector, 1 appears in the first entry. Then 1 diffuses from the first entry to the last entry as k increases, and the vector of all 1's is obtained when $k=2n$. For example, when $n=4$, the sequence of this type consists of $\boldsymbol{\alpha}^0=(1,1,1,1)$, $\boldsymbol{\alpha}^1=(-1,1,1,1)$, $\boldsymbol{\alpha}^2=(-1,-1,1,1)$, $\boldsymbol{\alpha}^3=(-1,-1,-1,1)$, $\boldsymbol{\alpha}^4=(-1,-1,-1,-1)$, $\boldsymbol{\alpha}^5=(1,-1,-1,-1)$, $\boldsymbol{\alpha}^6=(1,1,-1,-1)$, $\boldsymbol{\alpha}^7=(1,1,1,-1)$ and $\boldsymbol{\alpha}^8=(1,1,1,1)$.

The next proposition gives a formal statement of the result and a complete proof.

Proposition 1. *Let n be any integer greater than 1. Let $S=\{\boldsymbol{\alpha}^0,\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^{2n}\}$ be the sequence such that $\boldsymbol{\alpha}^k$ is given by*

$$\boldsymbol{\alpha}^k=\begin{cases}(\overbrace{-1,-1,\dots,-1}^{k},\overbrace{1,1,\dots,1}^{n-k}), & \text{if } 0\le k\le n,\\ (\underbrace{1,1,\dots,1}_{}\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\!\overbrace{}^{k-n},\overbrace{-1,-1,\dots,-1}^{2n-k}), & \text{if } n+1\le k\le 2n.\end{cases}$$

Then the RNN model described by (1)–(3) can generate S.

Proof. See Appendix A.1.

The sequence of the second type starts with the vector of all 1's. Next the first entry changes from 1 to -1. Then -1 moves from the first entry to the last entry as k increases, and the vector having -1 only in the last entry is obtained when $k=2n-1$. Finally the sequence returns to the vector of all 1's when $k=2n$. For example, when $n=4$, the sequence of this type consists of $\boldsymbol{\alpha}^0=(1,1,1,1)$, $\boldsymbol{\alpha}^1=(-1,1,1,1)$, $\boldsymbol{\alpha}^2=(-1,-1,1,1)$, $\boldsymbol{\alpha}^3=(1,-1,1,1)$, $\boldsymbol{\alpha}^4=(1,-1,-1,1)$, $\boldsymbol{\alpha}^5=(1,1,-1,1)$, $\boldsymbol{\alpha}^6=(1,1,-1,-1)$, $\boldsymbol{\alpha}^7=(1,1,1,-1)$ and $\boldsymbol{\alpha}^8=(1,1,1,1)$

The next proposition gives a formal statement of the result and a complete proof.

Proposition 2. *Let n be any integer greater than 1. Let $S=\{\boldsymbol{\alpha}^0,\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^{2n}\}$ be the sequence such that $\boldsymbol{\alpha}^k$ is given by*

$$\boldsymbol{\alpha}^k=\begin{cases}(1,1,\dots,1), & \text{if } k=0 \text{ or } k=2n,\\ (\overbrace{1,1,\dots,1}^{(k+1)/2-1},-1,\overbrace{1,1,\dots,1}^{n-(k+1)/2}), & \text{if } k \text{ is odd and } 1\le k\le 2n-1,\\ (\overbrace{1,1,\dots,1}^{k/2-1},-1,-1,\overbrace{1,1,\dots,1}^{n-k/2-1}), & \text{if } k \text{ is even and } 2\le k\le 2n-2.\end{cases}$$

Then the RNN model described by (1)–(3) can generate S.

Proof. See Appendix A.2.

If a sequence S can be generated by the RNN model described by (1)–(3), we immediately see that any sequence obtained from S by permuting entries of bipolar vectors. This is formally stated and proved as follows.

Proposition 3. *Let $S=\{\boldsymbol{\alpha}^0,\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^{m-1}\}$ be a sequence of n-dimensional bipolar vectors satisfying the conditions given in Theorem 1. If S can be generated by the RNN model described by (1)–(3) then, for any permutation matrix $\boldsymbol{P}$ of order n, the sequence $\tilde{S}=\{\boldsymbol{P}\boldsymbol{\alpha}^0,\boldsymbol{P}\boldsymbol{\alpha}^1,\dots,\boldsymbol{P}\boldsymbol{\alpha}^{m-1}\}$ can also be generated by the RNN model.*

Proof. Let $\boldsymbol{W}=\boldsymbol{W}^*$ and $\boldsymbol{b}=\boldsymbol{b}^*$ be the parameters of any RNN that generates the sequence S. We consider the RNN with the parameters $\boldsymbol{W}=\boldsymbol{P}\boldsymbol{W}^*\boldsymbol{P}^{\mathrm{T}}$ and $\boldsymbol{b}=\boldsymbol{P}\boldsymbol{b}^*$ where $\boldsymbol{P}$ is any permutation matrix of order n. The dynamics of this RNN is described by

$$\dot{\boldsymbol{x}}(t)=-\boldsymbol{x}(t)+\boldsymbol{P}\boldsymbol{W}^*\boldsymbol{P}^{\mathrm{T}}\boldsymbol{f}(\boldsymbol{x}(t))+\boldsymbol{P}\boldsymbol{b}^*.$$

Multiplying both sides by $\boldsymbol{P}^{\mathrm{T}}$ from left and setting $\tilde{\boldsymbol{x}}(t)=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{x}(t)$, we have

$$\dot{\tilde{\boldsymbol{x}}}(t)=-\tilde{\boldsymbol{x}}(t)+\boldsymbol{W}^*\boldsymbol{f}(\tilde{\boldsymbol{x}}(t))+\boldsymbol{b}^*.$$

Since this RNN generates S, all trajectories $\tilde{\phi}(t,\tilde{\boldsymbol{x}}^0)$ of the state vector $\tilde{\boldsymbol{x}}(t)$ such that $\tilde{\boldsymbol{x}}(0)=\tilde{\boldsymbol{x}}^0\in\mathcal{R}_{\boldsymbol{\alpha}^0}$ visits $\mathcal{R}_{(\boldsymbol{\alpha}^0+\boldsymbol{\alpha}^1)/2}$, $\mathcal{R}_{\boldsymbol{\alpha}^1}$, $\mathcal{R}_{(\boldsymbol{\alpha}^1+\boldsymbol{\alpha}^2)/2},\dots,\mathcal{R}_{(\boldsymbol{\alpha}^{m-2}+\boldsymbol{\alpha}^{m-1})/2}$ and $\mathcal{R}_{\boldsymbol{\alpha}^{m-1}}$. This means that all trajectories $\phi(t,\boldsymbol{x}^0)$ of the state vector $\boldsymbol{x}(t)$ such that $\boldsymbol{x}(0)=\boldsymbol{x}^0\in\mathcal{R}_{\boldsymbol{P}\boldsymbol{\alpha}^0}$ visits $\mathcal{R}_{(\boldsymbol{P}\boldsymbol{\alpha}^0+\boldsymbol{P}\boldsymbol{\alpha}^1)/2}$, $\mathcal{R}_{\boldsymbol{P}\boldsymbol{\alpha}^1}$, $\mathcal{R}_{(\boldsymbol{P}\boldsymbol{\alpha}^1+\boldsymbol{P}\boldsymbol{\alpha}^2)/2},\dots,\mathcal{R}_{(\boldsymbol{P}\boldsymbol{\alpha}^{m-2}+\boldsymbol{P}\boldsymbol{\alpha}^{m-1})/2}$ and $\mathcal{R}_{\boldsymbol{P}\boldsymbol{\alpha}^{m-1}}$. $\square$

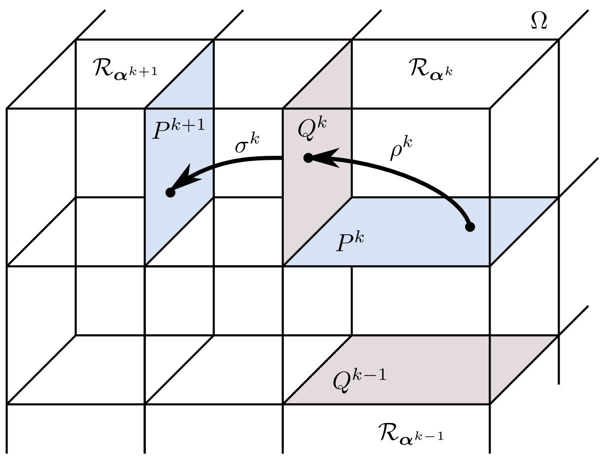


Fig. 8. Relationships among P^k, Q^k, P^{k+1}, ρ^k and σ^k in the case where $n = 3$.

6. Convergence of state trajectories of recurrent neural networks generating a periodic sequence

Suppose that, for a periodic sequence $S = \{\boldsymbol{\alpha}^0, \boldsymbol{\alpha}^1, \ldots, \boldsymbol{\alpha}^m (= \boldsymbol{\alpha}^0)\}$ of bipolar vectors, an RNN described by (1)–(3) satisfies (28)–(30) for $k = 0, 1, \ldots, m-1$. Then all trajectories $\phi(t, \boldsymbol{x}^0)$ of the state vector of the RNN such that $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}^0}$ visit $\mathcal{R}_{(\boldsymbol{\alpha}^0+\boldsymbol{\alpha}^1)/2}, \mathcal{R}_{\boldsymbol{\alpha}^1}, \mathcal{R}_{(\boldsymbol{\alpha}^1+\boldsymbol{\alpha}^2)/2}, \ldots, \mathcal{R}_{(\boldsymbol{\alpha}^{m-1}+\boldsymbol{\alpha}^m)/2}$ and $\mathcal{R}_{\boldsymbol{\alpha}^m} \left(= \mathcal{R}_{\boldsymbol{\alpha}^0}\right)$ repeatedly. However, it is not clear whether they converge to limit cycles or chaotic attractors. In this section, we analyze the behavior of such trajectories and derive a sufficient condition under which all of them converge to the same limit cycle. We also show through some examples that the sufficient condition is certainly satisfied.

In this section, we always use the following assumption.

Assumption 1. The sequence $S = \{\boldsymbol{\alpha}^0, \boldsymbol{\alpha}^1, \ldots, \boldsymbol{\alpha}^m\}$ satisfies the following conditions.

1. $\boldsymbol{\alpha}^{k+1}$ differs from $\boldsymbol{\alpha}^k$ in one and only one entry for $k = 0, 1, \ldots, m-1$.
2. $\boldsymbol{\alpha}^0, \boldsymbol{\alpha}^1, \ldots, \boldsymbol{\alpha}^{m-1}$ are different from each other, and $\boldsymbol{\alpha}^m = \boldsymbol{\alpha}^0$.
3. For each $i \in \{1, 2, \ldots, n\}$, there exists at least one $k \in \{0, 1, \ldots, m-1\}$ such that $\alpha_i^k \alpha_i^{k+1} = -1$.
4. There exist $\boldsymbol{W}$ and $\boldsymbol{b}$ that satisfy (28)–(30) for $k = 0, 1, \ldots, m-1$.

6.1. Positively invariant set

For an RNN described by (1)–(3), we define the sets $\Omega_i \subset \mathbb{R}^n$ for $i = 1, 2, \ldots, n$ by

$$\Omega_i := \{\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \mid l_i \le x_i \le u_i\}$$

where $l_i = -\sum_{j=1}^n |w_{ij}| + b_i$ and $u_i = \sum_{j=1}^n |w_{ij}| + b_i$. We further define the set $\Omega \subset \mathbb{R}^n$ by

$$\Omega := \Omega_1 \cap \Omega_2 \cap \cdots \cap \Omega_n = [l_1, u_1] \times [l_2, u_2] \times \cdots \times [l_n, u_n].$$

The next lemma shows that if an RNN generates S then the union of Ω and $\mathcal{R}_{\boldsymbol{v}}$, which defined in (4), is not empty for all $\boldsymbol{v} \in \{-1, 0, 1\}^n$.

Lemma 6. *If an RNN described by (1)–(3) satisfies (28)–(30) for $k = 0, 1, \ldots, m-1$ then $l_i < -1$ and $u_i > 1$ for $i = 1, 2, \ldots, n$.*

Proof. See Appendix A.3.

The next lemma shows that Ω is a positively invariant set for any RNN.

Lemma 7. *The set Ω is positively invariant for any RNN described by (1)–(3), that is, if $\boldsymbol{x}^0 \in \Omega$ then the trajectory $\phi(t, \boldsymbol{x}^0)$ of the state vector belongs to Ω for all $t > 0$. Moreover, if an RNN described by (1)–(3) satisfies (28)–(30) for $k = 0, 1, \ldots, m-1$ then for any $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}^0}$ there exists a $T \in [0, \infty)$ such that $\phi(T, \boldsymbol{x}^0) \in \Omega$.*

Proof. See Appendix A.4.

6.2. Convergence of state trajectories

In what follows, we focus our attention on RNNs described by (1)–(3) that satisfy (28)–(30) for $k = 0, 1, \ldots, m-1$. For such an RNN, we define the sets P^k and Q^k for $k = 0, 1, \ldots, m-1$ by

$$P^k := \partial\mathcal{R}_{(\boldsymbol{\alpha}^{(m-1+k) \bmod m}+\boldsymbol{\alpha}^k)/2} \cap \partial\mathcal{R}_{\boldsymbol{\alpha}^k} \cap \Omega,$$
$$Q^k := \partial\mathcal{R}_{\boldsymbol{\alpha}^k} \cap \partial\mathcal{R}_{(\boldsymbol{\alpha}^k+\boldsymbol{\alpha}^{k+1})/2} \cap \Omega.$$

We further define the mapping ρ^k from P^k to Q^k by $\rho^k(\boldsymbol{x}) := \phi(t^*, \boldsymbol{x})$ where $\boldsymbol{x}$ is a point in P^k and t^* is the smallest nonnegative number t satisfying $\phi(t, \boldsymbol{x}) \in Q^k$, and the mapping σ^k from Q^k to $P^{(k+1) \bmod m}$ by $\sigma^k(\boldsymbol{x}) := \phi(t^*, \boldsymbol{x})$ where $\boldsymbol{x}$ is a point in Q^k and t^* is the smallest positive number t satisfying $\phi(t, \boldsymbol{x}) \in P^{(k+1) \bmod m}$. Relationships among P^k, Q^k, P^{k+1}, ρ^k and σ^k are illustrated in Fig. 8. Using these notations, we define the mapping π from P^0 to itself as follows:

$$\pi := \sigma^{m-1} \circ \rho^{m-1} \circ \cdots \circ \sigma^0 \circ \rho^0.$$

For each $k \in \{0, 1, \ldots, m-1\}$, the Jacobian matrices of the mappings $\rho^k : P^k \to Q^k$ and $\sigma^k : Q^k \to P^{(k+1) \bmod m}$ are denoted by $\boldsymbol{J}_{\rho^k}$ and $\boldsymbol{J}_{\sigma^k}$, respectively. Also, the Jacobian matrix of the mapping $\pi : P^0 \to P^0$ is denoted by $\boldsymbol{J}_\pi$.

The next lemma guarantees the existence of a closed trajectory of the state vector passing through $P^0, Q^0, P^1, Q^1, \ldots, P^{m-1}, Q^{m-1}$.

Lemma 8. *There exists at least one point $\boldsymbol{x}^0 \in P^0$ such that $\phi(t, \boldsymbol{x}^0)$ forms a closed curve.*

Proof. It is clear that P^0 is a compact convex set in $\mathbb{R}^n$ and π is a continuous mapping from P^0 to itself. Therefore, by the Brouwer's fixed point theorem, there exists at least one point $\boldsymbol{x}^0 \in P^0$ such that $\pi(\boldsymbol{x}^0) = \boldsymbol{x}^0$ which means that $\phi(t, \boldsymbol{x}^0)$ forms a closed curve. □

In order to examine the convergence of the trajectories $\phi(t, \boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in P^0$ to the same limit cycle, we derive explicit formulas for the mappings ρ^k and σ^k and the matrix norm $\|\boldsymbol{J}_{\rho^k}\|_2$ and $\|\boldsymbol{J}_{\sigma^k}\|_2$ of their Jacobian matrices for $k = 0, 1, \ldots, m-1$, where $\|\cdot\|_2$ denotes the matrix norm induced by ℓ_2-norm of vectors. The results are shown in the next four lemmas.

Lemma 9. *For $k = 0, 1, \ldots, m-1$, the mapping $\rho^k : P^k \to Q^k$ is explicitly expressed as*

$$\rho_i^k(\boldsymbol{x}) = \frac{\left(\alpha_{i_k}^k - C_{i_k}^k\right)\left(x_i - C_i^k\right)}{x_{i_k} - C_{i_k}^k} + C_i^k, \quad i = 1, 2, \ldots, n \tag{42}$$

where

$$C_i^k := \sum_{j=1}^{n} w_{ij}\alpha_j^k + b_i,\ \ i = 1, 2, \ldots, n. \tag{43}$$

Proof. See Appendix A.5.

Lemma 10. *For $k = 0, 1, \ldots, m-1$, the matrix norm $\|\boldsymbol{J}_{\rho^k}\|_2$ is given by*

$$\|\boldsymbol{J}_{\rho^k}\|_2 = \begin{cases} \left|s_{i_k}^k\right|, & \text{if } n = 2,\\ \frac{1}{\sqrt{2}}\sqrt{\sum_{i=1}^{3}\left(s_i^k\right)^2 + \sqrt{\left(\sum_{i=1}^{3}\left(s_i^k\right)^2\right)^2 - 4\left(s_{i_k}^k s_{i_{k-1}}^k\right)^2}}, & \text{if } n = 3,\\ \max\left\{\left|s_{i_k}^k\right|, \frac{1}{\sqrt{2}}\sqrt{\sum_{i=1}^{n}\left(s_i^k\right)^2 + \sqrt{\left(\sum_{i=1}^{n}\left(s_i^k\right)^2\right)^2 - 4\left(s_{i_k}^k s_{i_{k-1}}^k\right)^2}}\right\}, & \text{otherwise} \end{cases} \tag{44}$$

where

$$s_i^k = s_i^k(\boldsymbol{x}) = \frac{\left(\alpha_{i_k}^k - C_{i_k}^k\right)\left(x_i - C_i^k\right)}{\left(x_{i_k} - C_{i_k}^k\right)^2},\ \ i = 1, 2, \ldots, n$$

with C_i^k being given by (43).

Proof. See Appendix A.6.

Lemma 11. *For $k = 0, 1, \ldots, m-1$, the mapping $\sigma^k : Q^k \to P^{(k+1) \bmod m}$ is explicitly expressed as*

$$\sigma_i^k(\boldsymbol{x}) = -\frac{\left(\alpha_{ik}^k + D_{i_k}^k\right)\left(x_i - D_i^k\right)}{\alpha_{i_k}^k - D_{i_k}^k} + D_i^k,\ \ i = 1, 2, \ldots, n \tag{45}$$

where

$$D_i^k := \sum_{j=1, j\neq i_k}^{n} w_{ij}\alpha_j^k + b_i,\ \ i = 1, 2, \ldots, n. \tag{46}$$

Proof. See Appendix A.7.

Lemma 12. *For $k = 0, 1, \ldots, m-1$, the matrix norm $\|\boldsymbol{J}_{\sigma^k}\|_2$ is given by*

$$\|\boldsymbol{J}_{\sigma^k}\|_2 = \left|\frac{\alpha_{i_k}^k + D_{i_k}^k}{\alpha_{i_k}^k - D_{i_k}^k}\right| \tag{47}$$

where D_i^k is given by (46).

Proof. See Appendix A.8.

We are now ready to present two results about the convergence of the trajectories $\phi(t, \boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}^0}$. The first one shows that when $n = 2$ the trajectories always converge to the same limit cycle. The second one provides a sufficient condition for the trajectories to converge to the same limit cycle. This sufficient condition can be tested algorithmically in a finite number of steps, while the condition provided in the previous work (Takahashi & Minetoma, 2008) requires an infinite number of steps.

Proposition 4. *If $n = 2$, all trajectories $\phi(t, \boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}^0}$ converge to the same limit cycle.*

Algorithm 1 Computation of an upper bound for $\|\boldsymbol{J}_{\pi}(\boldsymbol{x}^{0-})\|_2$

Input: A sequence $S = \{\boldsymbol{\alpha}^0, \boldsymbol{\alpha}^1, \ldots, \boldsymbol{\alpha}^m\}$ of $n(\geq 3)$ dimensional bipolar vectors satisfying Assumption 1 (let i_k be the unique index i such that $\alpha_i^k\alpha_i^{k+1} = -1$ for $k = 0, 1, \ldots, m-1$ and let i_{-1} be equal to i_{m-1} for convenience), a weight matrix $\boldsymbol{W} \in \mathbb{R}^{n\times n}$ and a bias vector $\boldsymbol{b} \in \mathbb{R}^n$ satisfying (28)–(30) for $k = 0, 1, \ldots, m-1$.

Output: An upper bound for $\|\boldsymbol{J}_{\pi}(\boldsymbol{x}^{0-})\|_2$ with $\boldsymbol{x}^{0-} \in P^0$

1: Set $\mathcal{B}^{0-} \leftarrow [l_1^{0-}, u_1^{0-}] \times [l_2^{0-}, u_2^{0-}] \times \cdots \times [l_n^{0-}, u_n^{0-}]$ where

$$[l_i^{0-}, u_i^{0-}] = \begin{cases} [\alpha_i^0, \alpha_i^0], & \text{if } \alpha_i^{m-1}\alpha_i^m = -1,\\ [1, \sum_{j=1}^{n}|w_{ij}| + b_i], & \text{if } \alpha_i^{m-1} = \alpha_i^m = 1,\\ [-\sum_{j=1}^{n}|w_{ij}| + b_i, -1], & \text{if } \alpha_i^{m-1} = \alpha_i^m = -1. \end{cases}$$

2: Set $k \leftarrow 0$.

3: Set $\bar{J}_\rho^k \leftarrow \frac{1}{\sqrt{2}}\sqrt{\sum_{i=1}^{3}\left(\bar{s}_i^k\right)^2 + \sqrt{\left(\sum_{i=1}^{3}\left(\bar{s}_i^k\right)^2\right)^2 - 4\left(\underline{s}_{i_k}^k \underline{s}_{i_{k-1}}^k\right)^2}}$ where

$$\bar{s}_i^k = \max_{\boldsymbol{x}\in\mathcal{B}^{k-}}\left\{\frac{(\alpha_{i_k}^k - C_{i_k}^k)(x_i - C_i^k)}{\left(x_{i_k} - C_{i_k}^k\right)^2}\right\},\ \ \underline{s}_i^k = \min_{\boldsymbol{x}\in\mathcal{B}^{k-}}\left\{\frac{(\alpha_{i_k}^k - C_{i_k}^k)(x_i - C_i^k)}{\left(x_{i_k} - C_{i_k}^k\right)^2}\right\}$$

and $C_i^k = \sum_{j=1}^{n} w_{ij}\alpha_j^k + b_i$ for $i = 1, 2, \ldots, n$.

4: If $n \geq 4$ and $\bar{s}_{i_k}^k > \bar{J}_\rho^k$ then set $\bar{J}_\rho^k \leftarrow \bar{s}_{i_k}^k$.

5: Set $\mathcal{B}^{k+} \leftarrow [l_1^{k+}, u_1^{k+}] \times [l_2^{k+}, u_2^{k+}] \times \cdots \times [l_n^{k+}, u_n^{k+}]$ where

$$l_i^{k+} = \min_{\boldsymbol{x}\in\mathcal{B}^{k-}}\left\{\frac{(\alpha_{i_k}^k - C_{i_k}^k)(x_i - C_i^k)}{x_{i_k} - C_{i_k}^k}\right\} + C_i^k,$$

$$u_i^{k+} = \max_{\boldsymbol{x}\in\mathcal{B}^{k-}}\left\{\frac{(\alpha_{i_k}^k - C_{i_k}^k)(x_i - C_i^k)}{x_{i_k} - C_{i_k}^k}\right\} + C_i^k$$

for $i = 1, 2, \ldots, n$.

6: Set $\bar{J}_\sigma^k \leftarrow \left|\frac{\alpha_{i_k}^k + D_{i_k}^k}{\alpha_{i_k}^k - D_{i_k}^k}\right|$ where $D_{i_k}^k = \sum_{j=1, j\neq i_k}^{n} w_{i_k j}\alpha_j^k + b_{i_k}$.

7: Set $\mathcal{B}^{(k+1)-} \leftarrow [l_1^{(k+1)-}, u_1^{(k+1)-}] \times [l_2^{(k+1)-}, u_2^{(k+1)-}] \times \cdots \times [l_n^{(k+1)-}, u_n^{(k+1)-}]$ where

$$l_i^{(k+1)-} = \min_{\boldsymbol{x}\in\mathcal{B}^{k+}}\left\{-\frac{(\alpha_{i_k}^k + D_{i_k}^k)(x_i - D_i^k)}{\alpha_{i_k}^k - D_{i_k}^k}\right\} + D_i^k,$$

$$u_i^{(k+1)-} = \max_{\boldsymbol{x}\in\mathcal{B}^{k+}}\left\{-\frac{(\alpha_{i_k}^k + D_{i_k}^k)(x_i - D_i^k)}{\alpha_{i_k}^k - D_{i_k}^k}\right\} + D_i^k$$

and $D_i^k = \sum_{j=1, j\neq i_k}^{n} w_{ij}\alpha_j^k + b_i$ for $i = 1, 2, \ldots, n$.

8: If $k = m-1$ then return $\prod_{k=0}^{m-1}\bar{J}_\rho^k\bar{J}_\sigma^k$ and stop. Otherwise set $k \leftarrow k+1$ and return to Step 3.

Proof. See Appendix A.9.

Proposition 5. *Let n be an integer greater than 2. If Algorithm 1 returns a number less than 1 then all trajectories $\phi(t, \boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}^0}$ converge to the same limit cycle.*

Proof. See Appendix A.10.

We finally give three examples to demonstrate the usefulness of the sufficient condition in Proposition 5. In each example, the parameters of the RNN model are determined from a given periodic sequence using the method to maximize the margin in the inequality constraints. Then the sufficient condition is tested

Table 1
The values of $\bar{J}_\rho^k$ and $\bar{J}_\sigma^k$ for $k = 0, 1, \ldots, m-1$ obtained by Algorithm 1 with the sequence S, the weight matrix $\boldsymbol{W}$ and the bias vector $\boldsymbol{b}$ given in Example 4.

k	0	1	2	3	4	5	6	7
$\bar{J}_\rho^k$	7.7978	1.8988	1.0521	1.3142	1.3124	1.3127	1.2434	0.9976
$\bar{J}_\sigma^k$	0.2000	0.1453	0.2000	0.2016	0.1453	0.2000	0.2016	0.2000

Table 2
The values of l_i^{m-} and u_i^{m-} for $i = 1, 2, \ldots, n$ obtained by Algorithm 1 with the sequence S, the weight matrix $\boldsymbol{W}$ and the bias vector $\boldsymbol{b}$ given in Example 4.

i	1	2	3	4
l_i^{m-}	1.7202028706	1.6554950975	2.5480797697	1
u_i^{m-}	1.7202031875	1.6554953742	2.5480800928	1

Table 3
The values of $\bar{J}_\rho^k$ and $\bar{J}_\sigma^k$ for $k = 0, 1, \ldots, m-1$ obtained by Algorithm 1 with the sequence S, the weight matrix $\boldsymbol{W}$ and the bias vector $\boldsymbol{b}$ given in Example 5.

k	0	1	2	3	4	5
$\bar{J}_\rho^k$	4.8153	2.6057	2.1177	1.9612	1.9277	1.9216
$\bar{J}_\sigma^k$	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111
k	6	7	8	9	10	11
$\bar{J}_\rho^k$	1.9206	1.9204	1.9204	1.9204	1.9204	1.9204
$\bar{J}_\sigma^k$	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111

by running Algorithm 1 with the sequence and the obtained parameters.

Example 4. Let S be the sequence of nine bipolar vectors considered in Example 2, and $\boldsymbol{W}$ and $\boldsymbol{b}$ be given by (40). Running Algorithm 1 with these inputs, we obtain

$$\prod_{k=0}^{7} \bar{J}_\rho^k \bar{J}_\sigma^k \approx 2.91853608 \times 10^{-4}$$

where the values of $\bar{J}_\rho^k$ and $\bar{J}_\sigma^k$ for $k = 0, 1, \ldots, 7$ are shown in Table 1. Since $\|\boldsymbol{J}_\pi(\boldsymbol{x}^0)\|_2$ is less than 1 for all $\boldsymbol{x}^0 \in P^0$, the mapping π has a unique fixed point $\boldsymbol{x}^* \in \mathcal{B}^{m-} = [l_1^{m-}, u_1^{m-}] \times [l_2^{m-}, u_2^{m-}] \times \cdots \times [l_n^{m-}, u_n^{m-}]$ where the values of l_i^{m-} and u_i^{m-} for $i = 1, 2, \ldots, n$ are shown in Table 2, and hence all trajectories $\phi(t, \boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in P^0$ converges to the same limit cycle formed by $\phi(t, \boldsymbol{x}^*)$.

Example 5. Let S be the sequence considered in Proposition 1 with $n = 6$ and

$$\boldsymbol{W} = \begin{bmatrix} 2.00 & -0.25 & -0.25 & -0.25 & -0.25 & -0.25 \\ 0.25 & 2.00 & -0.25 & -0.25 & -0.25 & -0.25 \\ 0.25 & 0.25 & 2.00 & -0.25 & -0.25 & -0.25 \\ 0.25 & 0.25 & 0.25 & 2.00 & -0.25 & -0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 & 2.00 & -0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0.25 & 2.00 \end{bmatrix}, \quad \boldsymbol{b} = \boldsymbol{0}.$$

The values of $\boldsymbol{W}$ and $\boldsymbol{b}$ are based on the proof of Proposition 1, and thus S, $\boldsymbol{W}$ and $\boldsymbol{b}$ satisfy the conditions (28)–(30). Running Algorithm 1 with S, $\boldsymbol{W}$ and $\boldsymbol{b}$ given above, we obtain

$$\prod_{k=0}^{13} \bar{J}_\rho^k \bar{J}_\sigma^k \approx 1.43752864 \times 10^{-7}$$

where the values of $\bar{J}_\rho^k$ and $\bar{J}_\sigma^k$ for $k = 0, 1, \ldots, 11$ are shown in Table 3. Since $\|\boldsymbol{J}_\pi(\boldsymbol{x}^0)\|_2$ is less than 1 for all $\boldsymbol{x}^0 \in P^0$, the mapping π has a unique fixed point $\boldsymbol{x}^* \in \mathcal{B}^{m-} = [l_1^{m-}, u_1^{m-}] \times [l_2^{m-}, u_2^{m-}] \times \cdots \times [l_n^{m-}, u_n^{m-}]$ where the values of l_i^{m-} and u_i^{m-} for $i = 1, 2, \ldots, n$ are shown in Table 4, and hence all trajectories $\phi(t, \boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in P^0$ converges to the same limit cycle formed by $\phi(t, \boldsymbol{x}^*)$.

Example 6. Let S be the sequence considered in Proposition 2 with $n = 6$ and

$$\boldsymbol{W} = \begin{bmatrix} 2.000 & -1.167 & -0.167 & -0.167 & -0.167 & -0.167 \\ 1.000 & 2.000 & -0.500 & 0.000 & 0.000 & 0.000 \\ -0.500 & 1.000 & 2.000 & -0.500 & 0.000 & 0.000 \\ 0.000 & -0.500 & 1.000 & 2.000 & -0.500 & 0.000 \\ 0.000 & 0.000 & -0.500 & 1.000 & 2.000 & -0.500 \\ 0.000 & 0.000 & 0.000 & -1.000 & 2.000 & 2.000 \end{bmatrix},$$

$$\boldsymbol{b} = \begin{bmatrix} 0.750 \\ 0.000 \\ 0.500 \\ 0.500 \\ 0.500 \\ 1.500 \end{bmatrix}.$$

The values of $\boldsymbol{W}$ and $\boldsymbol{b}$ are based on the proof of Proposition 2, and thus S, $\boldsymbol{W}$ and $\boldsymbol{b}$ satisfy the conditions (28)–(30). Running Algorithm 1 with S, $\boldsymbol{W}$ and $\boldsymbol{b}$ given above, we obtain

$$\prod_{k=0}^{11} \bar{J}_\rho^k \bar{J}_\sigma^k \approx 2.40378494 \times 10^{-6}$$

where the values of $\bar{J}_\rho^k$ and $\bar{J}_\sigma^k$ for $k = 0, 1, \ldots, 11$ are shown in Table 5. Since $\|\boldsymbol{J}_\pi(\boldsymbol{x}^0)\|_2$ is less than 1 for all $\boldsymbol{x}^0 \in P^0$, the mapping π has a unique fixed point $\boldsymbol{x}^* \in \mathcal{B}^{m-} = [l_1^{m-}, u_1^{m-}] \times [l_2^{m-}, u_2^{m-}] \times \cdots \times [l_n^{m-}, u_n^{m-}]$ where the values of l_i^{m-} and u_i^{m-} for $i = 1, 2, \ldots, n$ are shown in Table 6, and hence all trajectories $\phi(t, \boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in P^0$ converges to the same limit cycle formed by $\phi(t, \boldsymbol{x}^*)$.

In all the examples above, the output of Algorithm 1, an upper bound for $\|\boldsymbol{J}_\pi(\boldsymbol{x}^{0-})\|_2$, is much smaller than 1. Also, the authors have not found so far any RNN such that it generates a periodic sequence of bipolar vectors while Algorithm 1 returns a number greater than or equal to 1. This indicates the possibility that the trajectories $\phi(t, \boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in P^0$ always converge to the same limit cycle if the parameters of an RNN satisfy (28)–(30) for a periodic sequence of bipolar vectors.

7. Conclusions

We have studied the problem of finding the parameters of a continuous-time RNN model with a piecewise-linear activation function so that it generates a given sequence of bipolar vectors. First, we have derived a sufficient condition for the model to generate the given sequence under the assumption that all diagonal entries of the weight matrix are 2. This assumption makes the analysis tractable and allows us to obtain the sufficient condition in a simple form. A more general sufficient condition may be obtained if this assumption is removed, but it is not clear at this time what it would be. Second, we have proposed three methods based on mathematical programming to find the parameter values of the model that satisfy the sufficient condition, and observed that the method based on the maximization of the margin in the inequality constraints is better than others because the designed RNNs are not only robust against parameter variations but also able to output each bipolar vector for at least a certain period of time. Third, we have presented two types of sequences of bipolar vectors that can be generated by the model. Although only two types of sequences were considered in this paper, there may be other realizable sequences that are easy to characterize.

Table 4
The values of l_i^{m-} and u_i^{m-} for $i = 1, 2, \ldots, n$ obtained by Algorithm 1 with the sequence S, the weight matrix $\boldsymbol{W}$ and the bias vector $\boldsymbol{b}$ given in Example 5.

i	1	2	3	4	5	6
l_i^{m-}	1.0555411257	1.5553824164	2.0536366974	2.5344346961	2.8232226602	1
u_i^{m-}	1.0555411278	1.5553824187	2.0536367000	2.5344346998	2.8232226674	1

Table 5
The values of $\bar{J}_\rho^k$ and $\bar{J}_\sigma^k$ for $k = 0, 1, \ldots, m-1$ obtained by Algorithm 1 with the sequence S, the weight matrix $\boldsymbol{W}$ and the bias vector $\boldsymbol{b}$ given in Example 6.

k	0	1	2	3	4	5
$\bar{J}_\rho^k$	6.3771	1.2695	0.6558	2.2358	1.9560	2.0684
$\bar{J}_\sigma^k$	0.0408	0.2000	0.1107	0.2000	0.2000	0.2000
k	6	7	8	9	10	11
$\bar{J}_\rho^k$	2.2182	1.8658	2.5425	1.0325	1.9937	2.4784
$\bar{J}_\sigma^k$	0.2000	0.2000	0.2000	0.2000	0.2000	0.4286

Finally, focusing on the case where the model is designed so that it generates a periodic sequence of bipolar vectors, we have provided a sufficient condition for the trajectories of the state vector to converge to the same limit cycle. We also confirmed that this sufficient condition is satisfied for all the examples considered in this paper. This indicates the possibility that the trajectories always converge to the same limit cycle if we design an RNN using one of the proposed mathematical programming problems.

Future directions of this research are to derive a more general sufficient condition for the model to generate a given sequence of bipolar vectors without the assumption about the diagonal entries of the weight matrix, to develop new design procedures based on the general sufficient condition, to explore realizable sequences of bipolar vectors, and to conduct further analysis of the convergence of the trajectories of the state vector to a limit cycle.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request

Acknowledgments

The authors would like to thank anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper. This work was supported by JSPS KAKENHI, Japan Grant Number JP21H03510.

Appendix. Proofs

A.1. Proof of Proposition 1

As shown in Example 1, the sequence S for $n = 2$ is realizable. We thus assume hereafter that $n \geq 3$. Let the weight matrix $\boldsymbol{W}$ and the bias vector $\boldsymbol{b}$ be set as

$$w_{ij} = \begin{cases} 1/(n-2), & \text{if } j < i, \\ -1/(n-2), & \text{if } j > i, \end{cases} \tag{A.1}$$

and

$$\boldsymbol{b} = (0, 0, \ldots, 0). \tag{A.2}$$

The values of the left-hand side of (28)–(30) for the sequence S and the parameters given above are shown in Table A.7. For example, the left-hand side of the inequality (28) is given by

$$\begin{aligned} \alpha_{i_k}^k \left(\sum_{j=1, j \neq i_k}^{n} \alpha_j^k w_{i_k j} + b_{i_k} \right) &= -\sum_{j=1}^{i_k - 1} w_{i_k j} + \sum_{j=i_k+1}^{n} w_{i_k j} + b_{ik} \\ &= -\frac{1}{n-2} \cdot (i_k - 1) + \left(-\frac{1}{n-2} \right) \\ &\quad \times (n - i_k) + 0 \\ &= -\frac{n-1}{n-2}. \end{aligned}$$

which is less than -1. Also, it is easily seen from Table A.7 that the minimum value of the left-hand side of (29) and (30) is $-(n-3)/(n-2)$ and -1, respectively, for all i. Therefore the system of inequalities (28)–(30) is satisfied for all $n \geq 3$. □

A.2. Proof of Proposition 2

As shown in Example 1, the sequence S for $n = 2$ is realizable. We thus assume hereafter that $n \geq 3$. Let the weight matrix $\boldsymbol{W}$ and the bias vector $\boldsymbol{b}$ be set as

$$w_{ij} = \begin{cases} -1 - 1/n, & \text{if } i = 1 \text{ and } j = 2, \\ -1/n, & \text{if } i = 1 \text{ and } 3 \leq j \leq n, \\ 1, & \text{if } i = 2 \text{ and } j = 1, \\ -1/2, & \text{if } i = 2 \text{ and } j = 3, \\ 0, & \text{if } i = 2 \text{ and } 4 \leq j \leq n, \\ 0, & \text{if } 3 \leq i \leq n-1 \text{ and } 1 \leq j \leq i-3, \\ -1/2, & \text{if } 3 \leq i \leq n-1 \text{ and } j = i-2, \\ 1, & \text{if } 3 \leq i \leq n-1 \text{ and } j = i-1, \\ -1/2, & \text{if } 3 \leq i \leq n-1 \text{ and } j = i+1, \\ 0, & \text{if } 3 \leq i \leq n-1 \text{ and } i+2 \leq j \leq n, \\ 0, & \text{if } i = n \text{ and } 1 \leq j \leq n-3, \\ -1, & \text{if } i = n \text{ and } j = n-2, \\ 2, & \text{if } i = n \text{ and } j = n-1, \end{cases} \tag{A.3}$$

and

$$b_i = \begin{cases} 1 - 3/2n, & \text{if } i = 1, \\ 0, & \text{if } i = 2, \\ 1/2, & \text{if } 3 \leq i \leq n-1, \\ 3/2, & \text{if } i = n. \end{cases} \tag{A.4}$$

The values of the left-hand side of (28), (29) and (30) for the sequence S and the parameters given above are shown in Tables A.8–A.10, respectively. It is easy to see from these tables that the system of inequalities (28)–(30) is certainly satisfied for all $n \geq 3$. □

A.3. Proof of Lemma 6

We first prove that $l_i < -1$. It follows from Assumption 1 that, for each $i \in \{1, 2, \ldots, n\}$, there exists at least one $k \in$

Table 6
The values of l_i^{m-} and u_i^{m-} for $i = 1, 2, \dots, n$ obtained by Algorithm 1 with the sequence S, the weight matrix $\boldsymbol{W}$ and the bias vector $\boldsymbol{b}$ given in Example 6.

i	1	2	3	4	5	6
l_i^{m-}	1.2082127674	2.5000321964	2.5017550850	2.6831717299	2.4180760249	1
u_i^{m-}	1.2082127751	2.5000321975	2.5017550908	2.6831717537	2.4180760715	1

Table A.7
The values of the left-hand side (LHS) of (28)–(30) for the sequence S in Proposition 1 and the parameters $\boldsymbol{W}$ and $\boldsymbol{b}$ given in (A.1) and (A.2) when $n \geq 3$.

Eq.	Value of k	LHS
(28)	$1 \leq k \leq 2n-1$	$-(n-1)/(n-2)$
(29)	$1 \leq k \leq i-2$	$-(n-2(i-k)+1)/(n-2)$
	$i \leq k \leq n-1$	$(n+2(i-k)-1)/(n-2)$
	$n \leq k \leq n+i-2$	$(n+2(i-k)-1)/(n-2)$
	$n+i \leq k \leq 2n-1$	$(3n+2(i-k)-1)/(n-2)$
(30)	$1 \leq k \leq i-2$	$-(n-2(i-k)+2)/(n-2)$
	$i \leq k \leq n-1$	$(n+2(i-k)-2)/(n-2)$
	$n \leq k \leq n+i-2$	$(n+2(i-k)-2)/(n-2)$
	$n+i \leq k \leq 2n-1$	$(3n+2(i-k)-2)/(n-2)$

Table A.8
The values of the left-hand side (LHS) of (28) for the sequence S in Proposition 2 and the parameters $\boldsymbol{W}$ and $\boldsymbol{b}$ given in (A.3) and (A.4) when $n \geq 3$.

Value of k	LHS
$k = 0$	$-1-1/2n$
$k = 2$	$-1-3/2n$
$k = 1$ or $3 \leq k \leq 2n-2$	$-3/2$
$k = 2n-1$	$-5/2$

Table A.9
The values of the left-hand side (LHS) of (29) for the sequence S in Proposition 2 and the parameters $\boldsymbol{W}$ and $\boldsymbol{b}$ given in (A.3) and (A.4) when $n \geq 3$.

Value of i	Value of k	LHS
$i = 1$	$k = 1$	$1+1/2n$
	$k = 3$	$1+3/2n$
	$k = 4$	$1+7/2n$
	$k \in \{2\ell-1 \mid \ell = 3, 4, \dots, n\}$	$-1+3/2n$
	$k \in \{2\ell \mid \ell = 3, 4, \dots, n-1\}$	$-1+7/2n$
$2 \leq i \leq n-1$	$0 \leq k \leq \max\{0, 2i-7\}$ or $2i+3 \leq k \leq 2n-1$	$1/2$
	$k \in \{2i-6, 2i-5, 2i-2, 2i+1, 2i+2\} \cap \mathbb{Z}_{++}$	$3/2$
	$k \in \{2i-4, 2i-1\} \cap \mathbb{Z}_{++}$	$-1/2$
$i = n$	$0 \leq k \leq \max\{0, 2n-7\}$	$5/2$
	$k \in \{2n-6, 2n-5\} \cap \mathbb{Z}_{++}$	$9/2$
	$k = 2n-4$	$1/2$
	$k = 2n-2$	$3/2$

$\{0, 1, \dots, m-1\}$ such that $\alpha_i^k = -1$. So, for such a k, we have from (29) that

$$-\left(\sum_{j=1, j\neq i}^{n} w_{ij}\alpha_j^k + b_i\right) > -1.$$

The left-hand side of this inequality can be rewritten as

$$-\left(\sum_{j=1, j\neq i}^{n} w_{ij}\alpha_j^k + b_i\right) = -\left(\sum_{j=1}^{n} w_{ij}\alpha_j^k - w_{ii}\alpha_i^k + b_i\right) = -\left(\sum_{j=1}^{n} w_{ij}\alpha_j^k + 2 + b_i\right).$$

Table A.10
The values of the left-hand side (LHS) of (30) for the sequence S in Proposition 2 and the parameters $\boldsymbol{W}$ and $\boldsymbol{b}$ given in (A.3) and (A.4) when $n \geq 3$.

Value of i	Value of k	LHS
$i = 1$	$k = 1$	$-1/2n$
	$k = 3$	$1+5/2n$
	$k = 4$	$5/2n$
	$5 \leq k \leq 2n-2$	$-1+5/2n$
	$k = 2n-1$	$-1+1/2n$
$2 \leq i \leq n-1$	$k = 0$ (only if $i = 2$)	$-1/2$
	$0 \leq k \leq 2i-8$ or $2i+3 \leq k \leq 2n-1$ or $k \in \{2i-5, 2i-2\} \cap \mathbb{Z}_{++}$	$1/2$
	$k = 0$ (only if $i = 3$) or $k = 2i+1$ (only if $i = n-1$) or $k \in \{2i-7, 2i+2\} \cap \mathbb{Z}_{++}$	1
	$k \in \{2i-6\} \cap \mathbb{Z}_{++}$ or $k = 2i+1$ (only if $i \leq n-2$)	$3/2$
	$k \in \{2i-4, 2i-1\} \cap \mathbb{Z}_{++}$	-1
$k = n$	$0 \leq k \leq 2n-8$ or $k = 2n-5$	$5/2$
	$k = \max\{0, 2n-7\}$	$7/2$
	$k \in \{2n-6\} \cap \mathbb{Z}_{++}$	$9/2$
	$k \in \{2n-4, 2n-2\}$	$-1/2$

Hence we have

$$\sum_{j=1}^{n} w_{ij}\alpha_j^k + b_i < -1$$

which implies that $l_i < -1$. We next prove that $u_i > 1$. It follows from Assumption 1 that, for each $i \in \{1, 2, \dots, n\}$, there exists at least one $k \in \{0, 1, \dots, m-1\}$ such that $\alpha_i^k = 1$. So, for such a k, we have from (29) that

$$\sum_{j=1, j\neq i}^{n} w_{ij}\alpha_j^k + b_i > -1.$$

The left-hand side of this inequality can be rewritten as

$$\sum_{j=1, j\neq i}^{n} w_{ij}\alpha_j^k + b_i = \sum_{j=1}^{n} w_{ij}\alpha_j^k - w_{ii}\alpha_i^k + b_i = \sum_{j=1}^{n} w_{ij}\alpha_j^k - 2 + b_i.$$

Hence we have

$$\sum_{j=1}^{n} w_{ij}\alpha_j^k + b_i > 1$$

which implies that $u_i > 1$. □

A.4. Proof of Lemma 7

We first prove the first statement. Let i be any integer in $\{1, 2, \dots, n\}$. If $x_i = l_i$, we have

$$-x_i + \sum_{j=1}^{n} w_{ij} f_j(x_j) + b_i \geq -l_i - \sum_{j=1}^{n} |w_{ij}| + b_i = -l_i + l_i = 0.$$

Also, if $x_i = u_i$, we have

$$-x_i + \sum_{j=1}^{n} w_{ij} f_j(x_j) + b_i \le -u_i + \sum_{j=1}^{n} |w_{ij}| + b_i = -u_i + u_i = 0.$$

These inequalities mean that $-\boldsymbol{x} + \boldsymbol{W}\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{b}$ belongs to the tangent cone of the closed and convex set Ω_i for all $\boldsymbol{x} \in \partial\Omega_i$. Hence it follows from the Nagumo Theorem (Nagumo, 1942) that Ω_i is positively invariant. Therefore $\Omega = \Omega_1 \cap \Omega_2 \cap \cdots \cap \Omega_n$ is also positively invariant.

We next prove the second statement. Suppose that (28)–(30) hold for $k = 0, 1, \ldots, m-1$, and let $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}^0}$. Then it follows from the third condition of Assumption 1 that for each $i \in \{1, 2, \ldots, n\}$ there exists a $t_i \in [0, \infty)$ such that $\phi_i(t_i, \boldsymbol{x}^0) \in [-1, 1] \subset [l_i, u_i]$ or $\phi(t_i, \boldsymbol{x}^0) \in \Omega_i$. Moreover, because Ω_i is positively invariant, $\phi(t, \boldsymbol{x}^0) \in \Omega_i$ for all $t \ge t_i$. Therefore, if we set $t_{\max} := \max\{t_1, t_2, \ldots, t_n\} < \infty$, we have $\phi(t_{\max}, \boldsymbol{x}^0) \in \Omega$. □

A.5. Proof of Lemma 9

Let $\boldsymbol{x} \in P^k$. Since the dynamics of the RNN in $\mathcal{R}_{\boldsymbol{\alpha}^k}$ is expressed as

$$\dot{x}_i(t) = -x_i(t) + C_i^k, \; i = 1, 2, \ldots, n,$$

we have

$$\phi_i(t, \boldsymbol{x}) = \left(x_i - C_i^k\right) e^{-t} + C_i^k, \; 0 \le t \le t^*, \; i = 1, 2, \ldots, n$$

where t^* is the minimum nonnegative t such that $\phi(t, \boldsymbol{x}) \in Q^k$. Note that the mapping $\rho_i(\boldsymbol{x})$ is given by $\phi_i(t^*, \boldsymbol{x})$ and that t^* is the unique solution of the following equation:

$$\left(x_{i_k} - C_{i_k}^k\right) e^{-t} + C_{i_k}^k = \alpha_{i_k}^k.$$

Therefore, we have

$$\begin{aligned} \rho_i^k(\boldsymbol{x}) &= \phi_i(t^*, \boldsymbol{x}) \\ &= \left(x_i - C_i^k\right) e^{-t^*} + C_i^k \\ &= \left(x_i - C_i^k\right) \cdot \frac{\alpha_{i_k}^k - C_{i_k}^k}{x_{i_k} - C_{i_k}^k} + C_i^k, \; i = 1, 2, \ldots, n \end{aligned}$$

which completes the proof. □

A.6. Proof of Lemma 10

It follows from (42) that the (i, j)-th entry of $\boldsymbol{J}_{\rho^k}$ is given by

$$\left(\boldsymbol{J}_{\rho^k}\right)_{ij} = \frac{\partial \rho_i^k}{\partial x_j} = \begin{cases} s_{i_k}^k, & \text{if } i = j \text{ and } i \notin \{i_{k-1}, i_k\}, \\ -s_i^k, & \text{if } i \ne j \text{ and } j = i_k, \\ 0, & \text{otherwise.} \end{cases}$$

Also, the (i, j)-th entry of $(\boldsymbol{J}_{\rho^k})^{\mathrm{T}} \boldsymbol{J}_{\rho^k}$ is given by

$$\left((\boldsymbol{J}_{\rho^k})^{\mathrm{T}} \boldsymbol{J}_{\rho^k}\right)_{ij} = \begin{cases} \left(s_{i_k}^k\right)^2, & \text{if } i = j \notin \{i_k, i_{k-1}\}, \\ \sum_{i=1, i \ne i_k}^{n} \left(s_i^k\right)^2, & \text{if } i = j = i_k, \\ -s_{i_k}^k s_i^k, & \text{if } i \notin \{i_k, i_{k-1}\} \text{ and } j = i_k, \\ 0, & \text{otherwise.} \end{cases}$$

It is easy to see that the characteristic polynomial of $(\boldsymbol{J}_{\rho^k})^{\mathrm{T}} \boldsymbol{J}_{\rho^k}$ is expressed as

$$\begin{aligned} &\left|\lambda \boldsymbol{I} - (\boldsymbol{J}_{\rho^k})^{\mathrm{T}} \boldsymbol{J}_{\rho^k}\right| \\ &= \begin{cases} \lambda \left(\lambda - \left(s_{i_k}^k\right)^2\right), & \text{if } n = 2, \\ \lambda \left(\lambda - \left(s_{i_k}^k\right)^2\right)^{n-3} \left(\lambda^2 - \sum_{i=1}^{n} \left(s_i^k\right)^2 \lambda + \left(s_{i_k}^k s_{i_{k-1}}^k\right)^2\right), & \text{if } n \ge 3. \end{cases} \end{aligned}$$

Hence $(\boldsymbol{J}_{\rho^k})^{\mathrm{T}} \boldsymbol{J}_{\rho^k}$ has nonnegative eigenvalues 0 and $\left(s_{i_k}^k\right)^2$ if $n = 2$, and 0, $\left(s_{i_k}^k\right)^2$ (with algebraic multiplicity $n - 3$) and

$$\frac{1}{2}\left(\sum_{i=1}^{n} \left(s_i^k\right)^2 \pm \sqrt{\left(\sum_{i=1}^{n} \left(s_i^k\right)^2\right)^2 - 4\left(s_{i_k}^k s_{i_{k-1}}^k\right)^2}\right)$$

if $n \ge 3$. Since $\|\boldsymbol{J}_{\rho^k}\|_2$ is equal to the square root of the largest eigenvalue of $(\boldsymbol{J}_{\rho^k})^{\mathrm{T}} \boldsymbol{J}_{\rho^k}$, we have (44). □

A.7. Proof of Lemma 11

Let $\boldsymbol{x} \in Q^k$. Since the dynamics of the RNN in $\mathcal{R}_{(\boldsymbol{\alpha}^k + \boldsymbol{\alpha}^{k+1})/2}$ is expressed as

$$\dot{x}_i(t) = -x_i(t) + w_{ii_k} x_{i_k}(t) + D_i^k, \; i = 1, 2, \ldots, n,$$

we have

$$\phi_{i_k}(t, \boldsymbol{x}) = \left(\alpha_{i_k}^k + D_{i_k}^k\right) e^t - D_{i_k}^k, \; 0 \le t \le t^*$$

and

$$\begin{aligned} \phi_i(t, \boldsymbol{x}) = &\frac{1}{2} w_{ii_k} \left(\alpha_{i_k}^k + D_{i_k}^k\right) e^t + \left[x_i - \frac{1}{2} w_{ii_k} \left(\alpha_{i_k}^k - D_{i_k}^k\right) - D_i^k\right] e^{-t} \\ &- w_{ii_k} D_{i_k}^k + D_i^k, \; 0 \le t \le t^*, \; i \ne i_k, \end{aligned}$$

where t^* is the minimum nonnegative t such that $\phi(t, \boldsymbol{x}) \in P^{k+1}$. Note that the mapping $\sigma_i^k(\boldsymbol{x})$ is given by $\phi_i(t^*, \boldsymbol{x})$ and that t^* is the unique solution of the following equation:

$$\left(\alpha_{i_k}^k + D_{i_k}^k\right) e^t - D_{i_k}^k = -\alpha_{i_k}^k.$$

Therefore we have

$$\begin{aligned} \sigma_i^k(\boldsymbol{x}) &= \phi_i(t^*, \boldsymbol{x}) \\ &= \frac{1}{2} w_{ii_k} \left(\alpha_{i_k}^k + D_{i_k}^k\right) \left(-\frac{\alpha_{i_k}^k - D_{i_k}^k}{\alpha_{i_k}^k + D_{i_k}^k}\right) \\ &\quad + \left[x_i - \frac{1}{2} w_{ii_k} \left(\alpha_{i_k}^k - D_{i_k}^k\right) - D_i^k\right] \\ &\quad \times \left(-\frac{\alpha_{i_k}^k + D_{i_k}^k}{\alpha_{i_k}^k - D_{i_k}^k}\right) - w_{ii_k} D_{i_k}^k + D_i^k \\ &= \left(x_i - D_i^k\right) \left(-\frac{\alpha_{i_k}^k + D_{i_k}^k}{\alpha_{i_k}^k - D_{i_k}^k}\right) + D_i^k \end{aligned}$$

for $i \ne i_k$. This expression is valid also for $i = i_k$ because substituting $x_i = \alpha_{i_k}^k$ into it we obtain $-\alpha_{i_k}^k$. □

A.8. Proof of Lemma 12

It follows from (45) that the (i, j)-th entry of $\boldsymbol{J}_{\sigma^k}$ is given by

$$(\boldsymbol{J}_{\sigma^k})_{ij} = \frac{\partial \sigma_i^k}{\partial x_j} = \begin{cases} -\left(\alpha_{i_k}^k + D_{i_k}^k\right) / \left(\alpha_{i_k}^k - D_{i_k}^k\right), & \text{if } i = j \text{ and } i \ne i_k, \\ 0, & \text{otherwise.} \end{cases}$$

Hence $(\boldsymbol{J}_{\sigma^k})^{\mathrm{T}} \boldsymbol{J}_{\sigma^k}$ has nonnegative eigenvalues 0 and $\left(\alpha_{i_k}^k + D_{i_k}^k\right)^2 / \left(\alpha_{i_k}^k - D_{i_k}^k\right)^2$ (with algebraic multiplicity $n - 1$). Since $\|\boldsymbol{J}_{\sigma^k}\|_2$ is equal to the square root of the largest eigenvalue of $(\boldsymbol{J}_{\sigma^k})^{\mathrm{T}} \boldsymbol{J}_{\sigma^k}$, we have (47). □

A.9. Proof of Proposition 4

When $n = 2$, any sequence S satisfying the conditions in Theorem 1 and Assumption 1 consists of five bipolar vectors. So we assume hereafter that $m = 4$. Since all trajectories $\phi(t, \boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in \mathcal{R}_{\boldsymbol{\alpha}^0}$ reach P^0 within a finite period of time, it suffices for us to consider the case where $\boldsymbol{x}^0 \in P^0$. For each $\boldsymbol{x}^0 \in P^0$, the point at which the trajectory $\phi(t, \boldsymbol{x}^0)$ first returns to P^0 is given by $\pi(\boldsymbol{x}^0) = \sigma^3 \circ \rho^3 \circ \cdots \circ \sigma^0 \circ \rho^0(\boldsymbol{x}^0)$. Then we have

$$\left\|\boldsymbol{J}_\pi(\boldsymbol{x}^0)\right\|_2 = \left\|\boldsymbol{J}_{\sigma^3}\left(\boldsymbol{x}^{3+}\right)\boldsymbol{J}_{\rho^3}\left(\boldsymbol{x}^{3-}\right)\cdots\boldsymbol{J}_{\sigma^0}\left(\boldsymbol{x}^{0+}\right)\boldsymbol{J}_{\rho^0}(\boldsymbol{x}^0)\right\|_2$$

where $\boldsymbol{x}^{0+} = \rho^0(\boldsymbol{x}^{0-})$, $\boldsymbol{x}^{1-} = \sigma^0(\boldsymbol{x}^{0+})$, $\ldots$, $\boldsymbol{x}^{3-} = \sigma^2(\boldsymbol{x}^{2+})$ and $\boldsymbol{x}^{3+} = \rho^3(\boldsymbol{x}^{3-})$. From Lemmas 10 and 12 and the sub-multiplicativity of the matrix norm $\|\cdot\|_2$, we have

$$\begin{aligned}\left\|\boldsymbol{J}_\pi(\boldsymbol{x}^0)\right\|_2 &\le \left\|\boldsymbol{J}_{\sigma^3}\left(\boldsymbol{x}^{3+}\right)\right\|_2\left\|\boldsymbol{J}_{\rho^3}\left(\boldsymbol{x}^{3-}\right)\right\|_2\cdots\left\|\boldsymbol{J}_{\sigma^0}\left(\boldsymbol{x}^{0+}\right)\right\|_2\left\|\boldsymbol{J}_{\rho^0}(\boldsymbol{x}^0)\right\|_2 \\ &= \left|\frac{\alpha_{i_3}^3 + D_{i_3}^3}{\alpha_{i_3}^3 - D_{i_3}^3}\right|\left|\frac{\alpha_{i_3}^3 - C_{i_3}^3}{x_{i_3}^{3-} - C_{i_3}^3}\right|\cdots\left|\frac{\alpha_{i_0}^0 + D_{i_0}^0}{\alpha_{i_0}^0 - D_{i_0}^0}\right|\left|\frac{\alpha_{i_0}^0 - C_{i_0}^0}{x_{i_0}^0 - C_{i_0}^0}\right|.\end{aligned}$$

Here, it follows from (28) and (46) that

$$\begin{aligned}\left|\frac{\alpha_{i_k}^k + D_{i_k}^k}{\alpha_{i_k}^k - D_{i_k}^k}\right| &= \left|1 - \frac{2}{1 - \alpha_{i_k}^k D_{i_k}^k}\right| \\ &\le 1 - \frac{2}{1 - \min_{k\in\{0,1,2,3\}}\{\alpha_{i_k}^k D_{i_k}^k\}} \\ &< 1\end{aligned}$$

for $k = 0, 1, 2, 3$. Also, it follows from (28) and (43) that

$$\left|\frac{\alpha_{i_k}^k - C_{i_k}^k}{x_{i_k}^{k-} - C_{i_k}^k}\right| \le \left|\frac{\alpha_{i_k}^k - C_{i_k}^k}{\alpha_{i_k}^k - C_{i_k}^k}\right| = 1$$

for $k = 0, 1, 2, 3$. Thus we have

$$\left\|\boldsymbol{J}_\pi(\boldsymbol{x}^0)\right\|_2 \le \left(1 - \frac{2}{1 - \min_{k\in\{0,1,2,3\}}\{\alpha_{i_k}^k D_{i_k}^k\}}\right)^4 < 1$$

for all $\boldsymbol{x}^0 \in P^0$, which implies that π is a contraction mapping from P^0 to itself (Philips & Taylor, 1973). By the Banach fixed-point theorem, π has a unique fixed point $\boldsymbol{x}^* \in P^0$ and $\lim_{l\to\infty} \pi^l(\boldsymbol{x}^0) = \boldsymbol{x}^*$ for all $\boldsymbol{x}^0 \in P^0$. Therefore, the trajectory $\phi(t, \boldsymbol{x}^0)$ converges to the limit cycle formed by $\phi(t, \boldsymbol{x}^*)$ for all $\boldsymbol{x}^0 \in P^0$. □

A.10. Proof of Proposition 5

It follows from Theorem 1 and Lemma 7 that all trajectories $\phi(t, \boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in \mathcal{R}_{\alpha^0}$ reach $P^0 = \partial\mathcal{R}_{(\boldsymbol{\alpha}^{m-1}+\boldsymbol{\alpha}^0)/2} \cap \partial\mathcal{R}_{\alpha^0} \cap \Omega$ within a finite period of time. So we assume without loss of generality that $\boldsymbol{x}^0 \in P^0$. Algorithm 1 first computes $\mathcal{B}^{0-}$ which is equal to P^0, and then $\bar{J}_\rho^k$, $\mathcal{B}^{k+}$, $\bar{J}_\sigma^k$ and $\mathcal{B}^{(k+1)-}$ for $k = 0, 1, \ldots, m-1$. As for $\mathcal{B}^{k+}$ and $\mathcal{B}^{(k+1)-}$, we see from Lemmas 9 and 11 that they satisfy $\{\rho^k(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathcal{B}^{k-}\} \subseteq \mathcal{B}^{k+} \subseteq Q^k$ and $\{\sigma^k(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathcal{B}^{k+}\} \subseteq \mathcal{B}^{(k+1)-} \subseteq P^{k+1}$ for $k = 0, 1, \ldots, m-1$, which means that all trajectories $\phi(t, \boldsymbol{x}^0)$ with $\boldsymbol{x}^0 \in P^0$ pass through $\mathcal{B}^{0+}, \mathcal{B}^{1-}, \mathcal{B}^{1+}, \ldots, \mathcal{B}^{(m-1)-}, \mathcal{B}^{(m-1)+}$ and reach $\mathcal{B}^{m-} \subseteq \mathcal{B}^{0-} = P^0$. We thus can restrict the domains of the mappings ρ^k and σ^k to $\mathcal{B}^{k-}$ and $\mathcal{B}^{k+}$, respectively, for $k = 0, 1, \ldots, m-1$ to evaluate $\|\boldsymbol{J}_\pi(\boldsymbol{x}^0)\|_2$. Furthermore, we see from Lemmas 10 and 12 that $\bar{J}_\rho^k$ and $\bar{J}_\sigma^k$ are upper bounds for the matrix norm of the Jacobian matrices of the mappings $\rho^k\big|_{\mathcal{B}^{k-}} : \mathcal{B}^{k-} \to \mathcal{B}^{k+}$ and $\sigma^k\big|_{\mathcal{B}^{k+}} : \mathcal{B}^{k+} \to \mathcal{B}^{(k+1)-}$, respectively. Therefore

$$\|\boldsymbol{J}_\pi(\boldsymbol{x}^0)\|_2 \le \prod_{k=0}^{m-1} \bar{J}_\rho^k \bar{J}_\sigma^k$$

holds for all $\boldsymbol{x}^0 \in P^0$. If the right-hand side, which is the output of Algorithm 1, is less than 1 then we have $\|\boldsymbol{J}_\pi(\boldsymbol{x}^0)\|_2 < 1$ which implies that π is a contraction mapping from P^0 to itself (Philips & Taylor, 1973) and thus the trajectory $\phi(t, \boldsymbol{x}^0)$ converges to the same limit cycle. □

References

Aihara, K., Takabe, T., & Toyoda, M. (1990). Chaotic neural networks. *Physics Letters. A*, *144*(6–7), 333–340.

Amari, S. I. (1972). Learning patterns and pattern sequences by self-organizing nets of threshold elements. *IEEE Transactions on Computers*, *c-21*(11), 1197–1206.

Azhdari, M., & Binazadeh, T. (2021). Generation of stable limit cycles in nonlinear sandwich systems with dead-zone nonlinearity and input saturation. *Journal of the Franklin Institute*, *358*(18), 10029–10051.

Benmiloud, M., Benalia, A., Djemai, M., & Defoort, M. (2018). On the local stabilization of hybrid limit cycles in switched affine systems. *IEEE Transactions on Automatic Control*, *64*(2), 841–846.

Chen, X., Tao, Y., Xu, W., & Yau, S. S. T. (2022). Recurrent neural networks are universal approximators with stochastic inputs. *IEEE Transactions on Neural Networks and Learning Systems*.

Chua, L. O., & Yang, L. (1988a). Cellular neural networks: Theory. *IEEE Transactions on Circuits and Systems*, *35*(10), 1257–1272.

Chua, L. O., & Yang, L. (1988b). Cellular neural networks: Applications. *IEEE Transactions on Circuits and Systems*, *35*(10), 1273–1290.

Deng, K., Zhu, S., Bao, G., Fu, J., & Zeng, Z. (2023). Multistability of dynamic memristor delayed cellular neural networks with application to associative memories. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(2), 690–702.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179–211.

Funahashi, K. I., & Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks*, *6*(6), 801–806.

Hammer, B. (2000). On the approximation capability of recurrent neural networks. *Neurocomputing*, *31*(1–4), 107–123.

Hertrich, C., & Skutella, M. (2021). Provably good solutions to the knapsack problem via neural networks of bounded size. In *Proceedings of the AAAI conference on artificial intelligence, vol. 35, no. 9* (pp. 7685–7693).

Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*, *37*(1), 388–427.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, *79*(8), 2554–2558.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences*, *81*(10), 3088–3092.

Hopfield, J. J., & Tank, D. W. (1986). Computing with neural circuits: A model. *Science*, *233*(4764), 625–633.

Ijspeert, A. J. (2008). Central pattern generators for locomotion control in animals and robots: a review. *Neural Networks*, *21*(4), 642–653.

Jordan, M. I. (1997). Serial order: A parallel distributed processing approach. In *Advances in psychology, vol. 121* (pp. 471–495). Elsevier.

Kilian, J., & Siegelmann, H. T. (1996). The dynamic universality of sigmoidal neural networks. *Information and Computation*, *128*(1), 48–56.

Lawrence, S., Giles, C. L., & Fong, S. (2000). Natural language grammatical inference with recurrent neural networks. *IEEE Transactions on Knowledge and Data Engineering*, *12*(1), 126–140.

Lee, J. Y., & Dernoncourt, F. (2016). Sequential short-text classification with recurrent and convolutional neural networks. In *Proceedings of the 15th annual conference of the North American chapter of the association for computational linguistics: Human language technologies* (pp. 515–520).

Li, Z., Han, J., Weinan, E., & Li, Q. (2022). Approximation and optimization theory for linear continuous-time recurrent neural networks. *Journal of Machine Learning Research*, *23*, 1–85.

Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A critical review of recurrent neural networks for sequence learning. arXiv preprint arXiv:1506.00019.

Liu, P., Wang, J., & Zeng, Z. (2021). An overview of the stability analysis of recurrent neural networks with multiple equilibria. *IEEE Transactions on Neural Networks and Learning Systems*.

Michel, A. N., & Farrell, J. A. (1990). Associative memories via artificial neural networks. *IEEE Control Systems Magazine*, *10*(3), 6–17.

Mizuno, S. (1994). Polynomiality of infeasible-interior-point algorithms for linear programming. *Mathematical Programming*, *67*, 109–119.

Mori, Y., Davis, P., & Nara, S. (1989). Pattern retrieval in an asymmetric neural network with embedded limit cycles. *Journal of Physics A: Mathematical and General*, *22*(11), L525.

Nagumo, M. (1942). Über die lage der integralkurven gewöhnlicher differentialgleichungen. *Proceedings of the Physico-Mathematical Society of Japan*, *24*, 551–559.

Nie, X., & Zheng, W. X. (2015a). Multistability and instability of neural networks with discontinuous nonmonotonic piecewise linear activation functions. *IEEE Transactions on Neural Networks and Learning Systems*, *26*(11), 2901–2913.

Nie, X., & Zheng, W. X. (2015b). Multistability of neural networks with discontinuous non-monotonic piecewise linear activation functions and time-varying delays. *Neural Networks*, *65*, 65–79.

Nie, X., & Zheng, W. X. (2016). Dynamical behaviors of multiple equilibria in competitive neural networks with discontinuous nonmonotonic piecewise linear activation functions. *IEEE Transactions on Cybernetics*, *46*(3), 679–693.

Pan, P., Xu, Z., Yang, Y., Wu, F., & Zhuang, Y. (2016). Hierarchical recurrent neural encoder for video representation with application to captioning. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1029–1038).

Pearlmutter, B. A. (1989). Learning state space trajectories in recurrent neural networks. *Neural Computation*, *1*(2), 263–269.

Philips, G., & Taylor, P. (1973). *Theory and applications of numerical analysis*. New York: Academic Press.

Potra, F. A. (1996). An $O(nL)$ infeasible-interior-point algorithm for LCP with quadratic convergence. *Annals of Operations Research*, *62*(1), 81–102.

Sak, H., Senior, A. W., & Beaufays, F. (2014). Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In *INTERSPEECH* (pp. 338–342).

Salehinejad, H., Sankar, S., Barfett, J., Colak, E., & Valaee, S. (2017). Recent advances in recurrent neural networks. arXiv preprint arXiv:1801.01078.

Schäfer, A. M., & Zimmermann, H. G. (2007). Recurrent neural networks are universal approximators. *International Journal of Neural Systems*, *17*(04), 253–263.

Siegelmann, H. T., & Sontag, E. D. (1991). Turing computability with neural nets. *Applied Mathematics Letters*, *4*(6), 77–80.

Siegelmann, H. T., & Sontag, E. D. (1994). Analog computation via neural networks. *Theoretical Computer Science*, *131*(2), 331–360.

Siegelmann, H. T., & Sontag, E. D. (1995). On the computational power of neural nets. *Journal of Computer and System Sciences*, *50*(1), 132–150.

Takahashi, N., & Chua, L. O. (1998). On the complete stability of nonsymmetric cellular neural networks. *IEEE Transactions on Circuits and Systems I*, *45*(7), 754–758.

Takahashi, N., Ishitobi, K., & Nishi, T. (2010). Sufficient conditions for one-dimensional cellular neural networks to perform connected component detection. *Nonlinear Analysis. Real World Applications*, *11*(5), 4202–4213.

Takahashi, N., & Minetoma, Y. (2008). On asymptotic behavior of state trajectories of piecewise-linear recurrent neural networks generating periodic sequence of binary vectors. In *2008 IEEE international joint conference on neural networks* (pp. 484–489). IEEE.

Takahashi, N., Nagayoshi, M., Kawabata, S., & Nishi, T. (2008). Stable patterns realized by a class of one-dimensional two-layer CNNs. *IEEE Transactions on Circuits and Systems. I. Regular Papers*, *55*(11), 3607–3620.

Takahashi, N., Yamakawa, T., & Nishi, T. (2005). Realization of limit cycles by neural networks with piecewise linear activation function. In *Proceedings of the 2005 European conference on circuit theory and design, vol. 3* (pp. III/7–III/10). IEEE.

Williams, R. J., & Zipser, D. (1989). A learning algorithm for continually running fully recurrent neural networks. *Neural Computation*, *1*(2), 270–280.

Yoneyama, T., Ninomiya, H., & Asai, H. (2001). Design method of neural networks for limit cycle generator by linear programming. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, *84*(2), 688–692.

Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation*, *31*(7), 1235–1270.

Zhang, Z., Xu, Y., Yang, J., Li, X., & Zhang, D. (2015). A survey of sparse representation: algorithms and applications. *IEEE Access*, *3*, 490–530.